



STIC Search Report

Biotech-Chem Library

STIC Database Tracking Number: 119902

TO: **Phuong N Huynh**
Location: 3d75, 3c70
Saturday, April 24, 2004
Art Unit: 1644
Phone: 272-0846
Serial Number: 10 / 005907

From: Jan Delaval
Location: Biotech-Chem Library
Rem 1A51
Phone: 272-2504
jan.delaval@uspto.gov

Search Notes

Best Local Similarity: 85.19% Mismatches: 0
 Query Match: 81.68% Indels: 20
 DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK126682 (1-4032)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysProLys 20
 DB 377 ATGGGAATATCTCTCGGAATCTC-----403

QY 21 LysGlyAsnProAspGluGluArgLysValGlnGluMetThrThrPheGluArgLysLeu 40
 DB 404 -----AGGCAGGAATGACTACATTTGAAGAAACTT 436

QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAenGluAen 60
 DB 437 CAGAGAGATATGAAAGGCGAGATTCATCCACTTCTAATCAGGAACGAGAT 496

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 DB 497 GGCAGTCTATCTGAGGAAGTGTGTACACTCTATTATCATCATCCCATCCATCAGATCC 556

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAenIleAspSerLeuThrArgLysValArg 100
 DB 557 TCCCTGAGCTTCAATGATGATGGCTATGAGAACATTCATCCCTCCACAGGAAGTGA 616

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 DB 617 CAGTTAGAGAAAGTCTCAGACAGCAATATGCTTCTTAGACTTCTGTAGTAGGCTT 676

QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 677 TGTCTCTGACCCATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 721

RESULT 6
 AK124520 4136 bp mRNA linear PRI 09-SEP-2003
 LOCUS Homo sapiens cDNA FLJ43529 fis, clone BRACE3002390.
 DEFINITION AK124520
 VERSION AK124520.1 GI:34530324
 KEYWORDS oligo capping; fis (full insert sequence).
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T., Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K., Kamihara, K., Katsuta, N., Sato, K., Tanikawa, M., Yamazaki, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamae, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Murakawa, K., Kaneko, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Inoue, T.
 TITLE NEDO human cDNA sequencing project
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 4136)
 AUTHORS Isogai, T. and Yamamoto, J.
 TITLE Direct Submission
 JOURNAL Submitted (15-JUN-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Karusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
 (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
 COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing; Research Association for Biotechnology (RAB); cDNA library construction; Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5'- & 3'-end one pass sequencing; RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing; HRI and RAB; annotation; HRI and RAB.

FEATURES
 Location/Qualifiers
 1..4136

US-10-005-907-2 (1-135) x AK124520 (1-4136)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAenGlnLysProLys 20
 DB 498 ATGGGAATATCTCTCGGAATCTC-----524

QY 21 LysGlyAsnProAspGluGluArgLysValGlnGluMetThrThrPheGluArgLysLeu 40
 DB 525 -----AGGCAGGAATGACTACATTTGAAGAAACTT 557

QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAenGluAen 60
 DB 558 CAGATCATGATGAGAAAGGCGAGATTCATCCACTTCTAATCAGGAACGAGAT 617

QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 DB 618 GGCAGTGGTCTTGAAGAAAGTGTGTACACTCTATTATCATCCCATCCATCAGAAATCC 677

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAenIleAspSerLeuThrArgLysValArg 100
 DB 678 TCCCTGAGCTTCAATGATGATGGCTATGAGAACATTCATCCCTCCACAGGAAGTGA 737

QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 DB 738 CAGTTAGAGAAAGTCTCAGACAGCAATATGCTTCTTAGACTTCTGTAGTAGGCTT 797

QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 DB 798 TGTCTCTGACCCATGACATGATGATGATGATGATGATGATGATGATGATGATGAT 842

RESULT 7
 AC074365/c 141268 bp DNA linear HTG 23-SEP-2000
 LOCUS Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE.
 DEFINITION AC074365
 VERSION AC074365.5 GI:10280935
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (bases 1 to 141268)
 AUTHORS Waterston, R.H.
 TITLE The sequence of Homo sapiens clone
 JOURNAL Unpublished
 REFERENCE 2 (bases 1 to 141268)
 AUTHORS Waterston, R.H.

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at http://www.sanger.ac.uk/HGP/Chri RPI1-978115 is from the library RPI1-11.4 constructed by the group of Pieter de Jong. For further details see http://www.chori.org/bacpac/home.htm VECTOR: pBACe3.6

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

FEATURES

Location/Qualifiers
1..185467
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-978115"
/clone_lib="RPI1-11.4"

ORIGIN

Alignment Scores:
Pred. No.: 2,13e-33 Length: 185467
Score: 430.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.14% Indels: 0
Dbs: 9 Gaps: 0

US-10-005-907-2 (1-135) x AL606804 (1-185467)

QY 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysThrValLeuAsnHisle 75
Db 132080 CAGGAAACGAGATCGCAGTGGTCTGAAGAGTGTGCTACATGCTCAATACATC 132139
QY 76 ProHisGlnArgSerSerLeuSerSerAsnAspAspGlyThrGluAsnIleAspSerLeu 95
Db 132140 CCCATCAGAGATCTCCCTGAGCTCCATGATGATGGTATGAGAACATTCACCTCC 132199
QY 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluThrGluThrGluThrGluThr 115
Db 132200 ACAGGAAGTGTGACAGCTTTAGAGAAAGTGTGACAGAGATATGCTCTTCTAGGACT 132259
QY 116 SerValSerArgProCysSerCysThrHisGluHisAspThrGluValPheProHis 135
Db 132260 TCTGTTAGTAGGCTTGTCTCTGACCCATGACATGATGATGATGATGATGATGATGAT 132319

RESULT 9
BC030506
LOCUS
DEFINITION
Homo sapiens germinal center expressed transcript 2, mRNA (CDNA clone MGC:40441 IMAGE:4385178), complete cds.
1532 bp mRNA linear PRI 06-OCT-2003

ACCESSION
BC030506
VERSION
BC030506.1
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

REFERENCE
1 (bases 1 to 1532)
Srausberg,R.L., Feingold,R.A., Grouse,L.H., Derge,J.G., Klausner,R.D., Collins,F.S., Wagner,L., Shemen,C.M., Schuler,G.D., Altshuler,S.F., Zeeberg,K.H., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,K., Marusina,K., Parker,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Scheetz,T.E., Brownstein,M.J., Udwin,T.B., Toshiyuki,S., Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.J., Bosak,S.A., McEwan,P.J.,

McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,D.J., Hulyk,S.W., Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dikson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallos,D.E., Schnarch,A., Schein,J.E., Jones,S.J. and Marra,M.A. full-length Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1532)
Srausberg,R.
Direct Submission
Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: You Staudt
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC)
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc.mgc@nih.gov
Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Gupta,J., Haghighi,P., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masello,C., Maskari,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantrifop,S., Thomas,P.J., Touchman,J.W., Teague,A., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov Series: IRAC Plate: 64 Row: a Column: 15
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22749536.

FEATURES

Location/Qualifiers
1..1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40441 IMAGE:4385178"
/tissue_type="Lymph, lymphoma"
/clone_lib="NIH MGC 85"
/lab_host="DH10B"
/notes="vector: pCMV-SPORT6"
1..1532
/genes="GCET2"
/notes="synonyms: HGAL, GCAT2, MGC40441"
/db_xref="LocusID:257144"
/db_xref="MIN:607792"
54..590
/codon_start=1
/product="germinal center expressed transcript 2"
/protein_id="AAH30506.1"
/db_xref="GI:20987806"
/translation="MGNLLRENRQNTOMFWNVNQSPKORTSCWDHHAEGCFCLPKFKLIFSKQDSQNERMSSTPIQNVDTQSELCYTLINRVLCTPSPQMSABEYRNPVCKARPRSLGCTETETSLLEMPSTDPRHARSPEDVELLMPHRISSHP LQQRPLMAPSETQFSHL"

gene

CDS

ORIGIN

```

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
REFERENCE Bull, J.H., Ellison, G. and Paskins, L.D.
AUTHORS Diagnostic methods for the detection of prostate disorders
TITLE Patent: WO 0136674-A 20-MAY-2001;
JOURNAL AstraZeneca AB (SE)
FEATURES Location/Qualifiers
source 1..1933
            /organism="Homo sapiens"
            /mol_type="unassigned DNA"
            /db_xref="taxon:9606"
ORIGIN
Query Match 1.3%; Score 49; DB 6; Length 1933;
Best Local Similarity 100.0%; Fred.No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AATTCAGAAATGAAAAAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1878 AATTCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1926
RESULT 36
HUMSAPIA HUMSAPIA 1933 bp mRNA linear PRI 17-DEC-1993
LOCUS Homo sapiens SRP accessory protein 1A (SAP-1) mRNA, complete cds.
DEFINITION M85165
ACCESSION M85165
VERSION M85165.1 GI:429185
KEYWORDS serum response factor; SAP-1; Elk-1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1933)
Dalton.S. and Treisman,R.
Characterization of SAP-1, a protein recruited by serum response
factor to the c-fos serum response element
Cell 68 (3), 597-612 (1992)
JOURNAL MEDLINE 92154673
PUBMED 1339307
REFERENCE 2 (bases 1 to 1933)
Treisman,R.
Direct Submission
Submitted (05-MAR-1992) Richard Treisman, Transcription Laboratory,
Imperial Cancer Research Fund, London, England
On Nov 29, 1993 this sequence version replaced gi:338036.
COMMENT Original source text: Homo sapiens CDNA to mRNA.
Location/Qualifiers
1..1933
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
1..1933
    /gene="SAP-1"
    /gene="SAP-1"
    /note="Homology region A with Elk-1 protein (Ets domain)
is bp 150-417, amino acids 1-89; Homology region B with
Elk-1 protein, required for cooperative ternary complex
formation with SRF is bp 565-617, amino acids 136-157;
Sequence diverges from SAP-1B at bp 1229, amino acid 360;
Homology region C with Elk-1 protein, core of regulated
transcription activation domain, is bp 1203-1355, amino
acids 352-402; (SVT)P motifs conserved between SAP-1A and
Elk-1 are located at amino acids T354, T361-T366, S381,
S387, T420, S425, corresponding to bp 1209-1214, bp
1230-1235, bp 1245-1250, bp 1290-1295, bp 1308-1313, bp
1407-1412, bp-1422-1427"
/codon_start=1
/product="SAP-1A protein"
/product_id="FAAO3631.1"

```

FEATURES
Source Location/Qualifiers
1..527
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="from acute myelogenous leukemia"
/lab_hosts="DH10B (TI phage-resistant)"
/clone_lib="NH_MGC_55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgcttcggcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTAGTCC-3' and 3' adaptor sequence: 5'-ATTATAGAGCGGCGCGGACATG-dt(30)BN-3' (where B = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
Query Match 100.0%; Score 405; DB 10; Length 527;
Best Local Similarity 100.0%; Pred. No. 6.1e-86;
Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 60
DB 45 ATGGGAAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 104
QY 61 AAGGAAACCCAGATGAGAAAGAAACCGGAGGAAATGACTACATTTGAAAGAAACTT 120
DB 105 AAGGAAACCCAGATGAGAAAGAAACCGGAGGAAATGACTACATTTGAAAGAAACTT 164
QY 121 CAGCATCAAGATAGAAAGCCAGAGCTTCATCCACTTCTATCAGGAAACGAGAT 180
DB 165 CAGCATCAAGATAGAAAGCCAGAGCTTCATCCACTTCTATCAGGAAACGAGAT 224
QY 181 GGCAGTGGTTCTGAAGAGTGTGTACACTGTCTATTATCATCCGCCCATCAGAGATCC 240
DB 225 GGCAGTGGTTCTGAAGAGTGTGTACACTGTCTATTATCATCCGCCCATCAGAGATCC 284
QY 241 TCCCTGAGCTCCATGATGATGGCTTGAAGACATGATCCCTCCACAGGAAAGTGA 300
DB 285 TCCCTGAGCTCCATGATGATGGCTTGAAGACATGATCCCTCCACAGGAAAGTGA 344
QY 301 CAGTTAGAGAAAGTCAGAGACAGATATGCCCTTCTTAGGACTTCTGTAGAGCCT 360
DB 345 CAGTTAGAGAAAGTCAGAGACAGATATGCCCTTCTTAGGACTTCTGTAGAGCCT 404
QY 361 TGTTCCTGCCCCATGAGCATGATGAAAGTGTGTTCACAC 405
DB 405 TGTTCCTGCCCCATGAGCATGATGAAAGTGTGTTCACAC 449

RESULT 2
AW480906
LOCUS AW480906 288 bp mRNA linear EST 09-JUL-2000
DEFINITION 3516 MARC 2P1G Sus scrofa cDNA 5', mRNA sequence
ACCESSION AW480906
VERSION AW480906.1 GI:7050949
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (Bases 1 to 288)
REFERENCE 1 (Bases 1 to 288)
AUTHORS Fahrenkrug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wiss, A., Rohrer, G.A., Pertea, G., Sultana, R.,
Quackenbush, and Keele, J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly

Sat Apr 24 08:29:31 2004

us-10-005-907-1.oligo.rge

```
/db_xref="GI:429186"
/translation="MDSAITLWQFLQLLQKPNKHMICWTSNDGQFKLQAEVARL
WGIKKNPNMYDKLSALRYVYVKNIIKKVNGKFFVYVYXPEILNMDPMTVGRIE
GDCESLNFSVSSXDVNGGKPPQPGAKTSRNDYIHSGLYSSFTLNSINSNV
KLYKLTENPAEKLAEEKPQBPSTSVIKFVTPSKKPPVEPVATISIGSISPS
BETIOALETVASPKLSELPASASNMVFAATTPRISSIDPLQEPPTSPFLSHHP
DIDTIDSVASQPMELFENLSLEPKQDSVLEKKNVNSRSKPKGGLGAPLVIT
SSDEPLGLSPSLPTASLTPAFFSQTPILUTPSELLSSIHFWSTLSFVAPLSAPLQ
CANTLFOFFSVLNSHGPTTSLGMDGPESTEGEPFDLQKT"
```

ORIGIN

```
Query Match 1.8% Score 49; DB 9; Length 1933;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1878 AATGCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1926
```

BLASTN-67

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 13:36:26 ; Search time 2525 Seconds
(without alignments)
4789.775 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405

Sequence: 1 atgggaattattctctg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_prt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	100.0	527	10	BF242113 601880401
2	64.8	16.0	288	10	AM480906 33515 MAR
3	58.6	14.7	1045	10	BF797507 602257174
4	58.6	14.5	1085	12	EM455407 AGENCOURT

5	58	14.3	508	29	AY404586
6	58	14.3	876	13	BQ214406
7	58	14.3	924	12	BG389415
8	58	14.3	987	9	AL560376
9	58	14.3	1000	12	BM904106
10	58	14.3	1039	12	BM479887
11	58	14.3	1070	12	BM456595
12	58	14.3	1095	12	BG106563
13	58	14.3	1124	10	BF795688
14	58	14.3	1193	12	BM455198
15	57.4	14.2	526	14	CB286394
16	56.4	13.9	683	10	AM963366
17	56.2	13.9	1165	12	BM458339
18	48.4	12.0	495	9	AA311038
19	47.2	11.7	1041	10	BG024891
20	47	11.6	1059	12	BM459235
21	46.8	11.6	353	29	AY404587
22	46.6	11.5	259	29	CE192340
23	45.4	11.2	658	29	CE020416
24	45	11.1	967	14	CD105480
25	44.8	11.1	826	13	BU746391
26	44.8	11.1	840	13	BU746392
27	44.6	11.0	560	13	BU695543
28	44.6	11.0	774	28	AI9721
29	44.4	11.0	731	29	CE188479
30	44	10.9	203	29	CE091328
31	44	10.9	421	28	AQ501175
32	44	10.9	520	12	BG145241
33	44	10.9	522	12	BG149093
34	44	10.9	524	12	BG276802
35	44	10.9	526	12	BG276546
36	44	10.9	668	29	CE207791
37	44	10.9	825	14	CB320849
38	43.8	10.8	899	29	CNS030Q4
39	43.8	10.8	959	29	CNS00655
40	43.8	10.8	1201	9	AL514529
41	43.6	10.8	815	13	BU932945
42	43.4	10.7	404	13	EX367419
43	42.8	10.6	703	29	CE790536
44	42.8	10.6	756	29	CE196221
45	42.6	10.5	891	10	BF264660

ALIGNMENTS

RESULT 1
BF242113
LOCUS 601880401F1.MH.MGC_55 Homo sapiens cDNA clone IMAGE:4109064 5',
DEFINITION 527 bp mRNA linear EST 14-NOV-2000
ACCESSION BF242113.1 GI:11156040
VERSION BF242113.1
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens (human)
REFERENCE 1 (bases 1 to 527)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LCM993 row: a column: 01
High quality sequence stop: 518.

JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE 22213789
PUBMED 12226715
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 20
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACAT
BACKWARD: GTTTCACGTCACGACG
Plate: 19 row: F column: 21
Seq primer: ATTAGTGACACTAG.
Location/Qualifiers
1. .288
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon.9823"
/tissue_type="pooled"
/lab_host="DH10B"
/clone_lib="MARC 2P1G"
/notes="Vector: pCMV SPORT6; Site.1: NotI; Site.2: SalI;
Library made from pooled tissue from testis, ovary,
endometrium, hypothalamus, pituitary, and placenta."

ORIGIN
Query Match 16.0%; Score 64.8; DB 10; Length 288;
Best Local Similarity 75.8%; Pred. No. 4e-05;
Matches 94; Conservative 0; Mismatches 27; Indels 3; Gaps 1;
QY 285 CACAAGAAAGTCAGACAGCTTTAGAGAAAGCTCAGACACAGAAATATGCCCTTCTTAGGAC 344
DB 2 CACAAGAGAGTCAGACCATATAGAGAGAGATCAGAAACAGATATGCCCTTCTCAGGAC 61
QY 345 TTCTGTAGT---AGGCTTGTCTGCACCCATGACGATGATATGAAGTTGTGTTCC 401
DB 62 GACTTACATCCAAAGTTTTCCTCTATACCCCTGAGATGATTGAAGTGTGTCTCC 121
QY 402 ACAC 405
DB 122 TCAC 125

RESULT 3
BF797507 1045 bp mRNA linear EST 12-JAN-2001
LOCUS 60225717AF1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4340323 5',
DEFINITION mRNA sequence.
ACCESSION BF797507
VERSION BF797507.1 GI:12102561
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1045)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: rcspbs@email.nih.gov
Tissue Procurement: Louis Staudt, M.D.; Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LMAN952 row: d column: 20

High quality sequence stop: 687.

FEATURES
Location/Qualifiers
1..1045

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:4340323"
/tissue_type="lymphoma, cell line"
/lab_hosts="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 14.7%; Score 59.6; DB 10; Length 1045;
Best Local Similarity 54.7%; Pred. No. 0.00075;
Matches 163; Conservative 0; Mismatches 129; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCAGGAAAA 173
DB 267 AAAGAGGACAGATTCCCAAAACCAAAATGAAGAATGTCTACTTCCCATCCAGGACAA 326

QY 174 CGAGATGGCAGTGGTCTGAGAGAGTGCTACCTGCTCATTAATCACA---TCCCCCA 230
DB 327 TGTGACACAGACTACTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTCTCTG 386

QY 231 TCAGAGATCCCTCCTGAGCTCCAAATGATGAGTATGAGAACTGATGCTCCCTCACAAG 290
DB 387 TACAAGCCATCAGGGAACCTGCTGAGAGAGTACTATGAGATGTTTCCCTGCAAGCTGA 446

QY 291 GAAAGTGAGACAGTTAGAGAAAGTCCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
DB 447 GAGACCCAGAGTCTCTGAGAGAACTGAGACTGAGTATGCTTCTACATATGCTTTC 506

QY 348 TGTAGTAGGCTTGTCTGACCCATGAGCATGATTAAGATGTTGTTTCCACAC 405
DB 507 TACAGACCCAGGATGCCGATCCCAAGAGATGATATGAATGAACTTCTCATGCTCAC 564

RESULT 4
LOCUS BM455407 1085 bp mRNA linear EST 05-FEB-2002
DEFINITION AGENCOURT_6407334 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500391
ACCESSION BM455407
VERSION BM455407.1 GI:18504436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Straubeberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM12135 row: d column: 24
High quality sequence start: 19
High quality sequence stop: 760.
Location/Qualifiers
1..1085
/organism="Homo sapiens"

FEATURES
source

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:5500391"

/tissue_type="lymphoma, cell line"

/lab_hosts="DH10B (phage-resistant)"

/clone_lib="NIH_MGC_85"

/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI; Site: 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 14.5%; Score 58.6; DB 12; Length 1085;
Best Local Similarity 54.4%; Pred. No. 0.0013;
Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCAGGAAAA 173
DB 459 AAAGAGGACAGATTCCCAAAACCAAAATGAAGAATGTCTACTTCCCATCCAGGACAA 518

QY 174 CGAGATGGCAGTGGTCTGAGAGAGTGCTACACTGTCTAATTAATCACA---TCCCCCA 230
DB 519 TGTGACACAGACTACTCAGAGAGCTGTGCTATACCTCATCATCATCGGGTCTCTG 578

QY 231 TCAGAGATCCCTCCTGAGCTCCAAATGATGAGTATGAGAACTGATGCTCCCTCACAAG 290
DB 579 TACAAGCCATCAGGGAACCTGCTGAGAGAGTACTATGAGATGTTTCCCTGCAAGCTGA 638

QY 291 GAAAGTGAGACAGTTAGAGAAAGTCCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
DB 639 GAGACCCAGAGTCTCTGAGAGAACTGAGACTGAGTATGCTTCTACATATGCTTTC 698

QY 348 TGTAGTAGGCTTGTCTGACCCATGAGCATGATTAAGATGTTGTTTCCACAC 405
DB 699 TACAGACCCAGGATGCCGATCCCAAGAGATGATATGAATGAACTTCTCATGCTCAC 756

RESULT 5
LOCUS AY404586 508 bp DNA linear GSS 16-DEC-2003
DEFINITION Homo sapiens GCET2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY404586
VERSION AY404586.1 GI:39760563
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 508)
AUTHORS Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A., Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B., Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Shinsky, J.J., Adams, M.D. and Cargill, M.
TITLE Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.
FEATURES
source
1..508
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db xref="taxon:9606"
 <1..>508
 /gene="GCEr2"
 /locus_tag="HCM1940"

ORIGIN

Query Match 14.3%; Score 58; DB 29; Length 508;
 Best Local Similarity 54.4%; Pred. No. 0.0017;
 Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCAGAGCTTCATCCACTTCAATCAGGAAA 173
 DB 136 AAAGAGCAGATTCCTCAAAACGAAATGAAGATGCACTCTCCATCCAGGAAA 195
 QY 174 CGAAGTGGCAGTGGTTCGAGAGTGGCTACACTGTCATTAATACACA---TCCGCCA 230
 DB 196 TGTGACCAACCTACTCAGAGAGCTGTGCTATACCTCATCAATCATGGGTCTCTG 255
 QY 231 TCAGAGATCTCCCTCAGCTCCAAATGATGCTGATGAGAACTGACTCCCTCAAG 290
 DB 256 TACAGGCCATCAGGAACTCTGCTGAGAGTACTATGAGATGTTCCCTGCAAGCTGA 315
 QY 291 GAAAGTGAACATTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCT---TAGGACTTC 347
 DB 316 GAGACCCAGAGATCTCTGGAGAACTGAGACTGAGTATTCACTTCTACATATGCCCTTC 375
 QY 348 TGTAGTAGGCTTGTCTCTGACCCATGAGCATGATTATGAAGTGTGTTTCCACAC 405
 DB 376 TACAGACCCAGGATGCGCGATCCCGAGAGATGAATGATGAATCTCATGCGCTCAC 433

RESULT 6
 BQ214406
 LOCUS
 DEFINITION AGENCOURT_7592611 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6070069
 5', mRNA sequence.
 ACCESSION BQ214406
 VERSION BQ214406.1 GI:20395806
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 876)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLMI3353 row: m column: 14
 High quality sequence stop: 678.
 Location/Qualifiers
 1..876
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:6070069"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally, oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN

Query Match 14.3%; Score 58; DB 13; Length 876;
 Best Local Similarity 54.4%; Pred. No. 0.0018;
 Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAAATTCAGATCAAGATGAAGAAAGCCAGAGCTTCATCCACTTCAATCAGGAAA 173
 DB 404 AAAGAGCAGATTCCTCAAAACGAAATGAAGATGCACTCTCCATCCAGGAAA 463
 QY 174 CGAAGTGGCAGTGGTTCGAGAGTGGCTACACTGTCATTAATACACA---TCCGCCA 230
 DB 464 TGTGACCAACCTACTCAGAGAGCTGTGCTATACCTCATCAATCATGGGTCTCTG 523
 QY 231 TCAGAGATCTCCCTCAGCTCCAAATGATGCTGATGAGAACTGACTCCCTCAAG 290
 DB 524 TACAAGGCCATCAGGAACTCTGCTGAGAGTACTATGAGATGTTCCCTGCAAGCTGA 583
 QY 291 GAAAGTGAACATTTAGAGAAAGGTCAGAGACAGAAATATGCCCTTCT---TAGGACTTC 347
 DB 584 GAGACCCAGAGATCTCTGGAGAACTGAGACTGAGTATTCACTTCTACATATGCCCTTC 643
 QY 348 TGTAGTAGGCTTGTCTCTGACCCATGAGCATGATTATGAAGTGTGTTTCCACAC 405
 DB 644 TACAGACCCAGGATGCGCGATCCCGAGAGATGAATGATGAATCTCATGCGCTCAC 701

RESULT 7
 BQ389415
 LOCUS
 DEFINITION 602414106F1 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:4522512 5', mRNA sequence.
 ACCESSION BQ389415
 VERSION BQ389415.1 GI:13282861
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 NIH-MGC http://mgi.nci.nih.gov/
 1 (bases 1 to 924)
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLMI0423 row: l column: 01
 High quality sequence stop: 726.
 Location/Qualifiers
 1..924
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4522512"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally, oligo-dT primed. Average insert size 2.5 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

FEATURES

source

QY	174	CGAGAAATGGCAGTGGTTCTGGAAGAAGTGCTCACTGTGCATTAAATCACAAA---TCCCCCA	230
DB	403	TGTTGACCAGACCTACTTCAGAGAGGTGTGCTATACCTCATCAATCATCGGGTTCCTGTG	462
QY	231	TCAGAGATCCCTCCCTGAGCTCCCAATGATGCTATGAGAATGCTCTCCTCACAAG	290
DB	463	TACAAGGCCATCGGGAACCTCTCTGAAGAGTACTATGAGATGTTCCCTGCAGAACTGA	522
QY	291	GAAAGTGAGACAGATTTAGAGAAAGGTGAGACAGAAATATGCCCTTCT---	347
DB	523	GAGACCCAGAGAGTCTTGGGAGAACTGAGACTGAGTATTCACTTCTACATATGCTTTC	582
QY	348	TGTTAGTAGGGCCTTGTTCCCTGCACCCATGACGACGATTAAGAATGTGTGTTTCCACAC	405
DB	583	TACAGACCCCAGGCATGCCGATCCCGATCCCAAGAGATGAATGAATCTCATGSCCTCAC	640

BM904106	1000 bp	mRNA	linear	EST 12-MAR-2002			
AGENCOURT_6692745	NIH_MGC_85	Homo sapiens	cDNA clone	IMAGE:5499556			
5'	mRNA sequence.						
BM904106							
BM904106.1	GI:19354351						
EST.							
Homo sapiens	(human)						
Homo sapiens							
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.							
1 (Bases 1 to 1000)							
NIH-MGC	http://mgc.ncbi.nih.gov/ .						
National Institutes of Health, Mammalian Gene Collection (MGC)							
Unpublished (1999)							
Contact: Robert Strausberg, Ph.D.							
Email: cgapbs-x@mail.nih.gov							
Tissue Procurement: Lou Staudt							
cDNA Library Preparation: Life Technologies, Inc.							
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)							
DNA Sequencing by: Agencourt Bioscience Corporation							
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:							
http://image.llnl.gov							
Plate: LLAM12133	row: b	column: 05					
High quality sequence stop: 698.							

```

CDNA library Arrayed by: INS I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12133 row: b column: 05
High quality sequence stop: 698
Location/Qualifiers
1. 1000
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5499556"
/tissue_type="lymphoma, cell line"
/lab_host="DH105 (phage-resistant)"
/clone_lib="NIH MGC_95"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI;
Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."
FEATURES
source

```

au_xcel="taxid:3808"
 /clone="IMAGE:549556"
 /tissue_type="lymphoma, cell line"
 /lab_host="DHIOB (phage-resistant)"
 /clone_lib="NIH MGC_95"
 /notes="Organ: lymph; Vector: pCMV-SPORT6; Site: 1: NotI;
 Site: 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH MGC Library."

Matches	162;	Conservative	0;	Mismatches	130;	Indels	6;	Gaps
QY	114	AAAATCTAAGATCAAGATAGAAAAGCCAGAAGTTTCATCCACTCTCTAATTCAGGAAA	173					
Db	281	AANGAGGCAAGATGCCAAAACGAAAAATGAAGAATGTCATCTACTCCCATCAGAGCAA	340					
QY	174	CGAAGATGGCAGTGGTTCTGAAGAAAGTGCTGACACTCTCAATTAAACACA---	230					
Db	341	TGTTGACACAGACCTACTCAGAGGAGCTGTGTATACCCCTCATCAATCATGGGTTCTCTG	400					

Qy	231	TCAGAGATCTCCCTGAGCTCCATGATGATGGCTATGAGAAATATGATCCCTCAAG	290
Db	401	TACAAGGCCATCAGGGAACTCTGCTGAAGAGTACTATGAGAAATGTTCCCTCAAAAGCTGA	460
Qy	291	GAAGTGAAGACAGTTTATAGAAAAGTCTAGACAGACATATGCCCTTCT---	347
Db	461	GAGACCCAGAGAGTCTTGGAGAGAACTGAGACTGAGTATCACTTCTACATATAGCCTTC	520
Qy	348	TGTTAGTAGGCTTGTTCCTCCACCCCATGAGCATGATTTATGAAGTGTGTTTCCACAC	405
Db	521	TACAGACCCAGGAGTCCCGATCCCAAGAGATGAATATGAATCTTCTATGCGCTCAC	578

[illegible]

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/ILNL at:

```

step://img2.cim-90v
Place: LHAM12330 row: 1 column: 05
High quality sequence stop: 665.
Location/Qualifiers
1. 1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5577364"
/tissue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 92"
/notes="Organ: testis; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH MGC Library."

```

Query Match	14.3%	Score 58;	DB 12;	Length 1039;
Best Local Similarity	54.4%;	Pred. No.	0.0018;	
Matches 162;	Conservative	0;	Mismatches 130;	Indels 6;
				Gaps 2

114	QY	AAAACTTCAGATCAAGATAGAAAAGCCAGAAAGTTTCATCCATCTTAATCAGGAAA	173
245	Db	AAAGAGGCAAGATCCCAAACGAAAACGAAAGATGTCACTACTCCCATCCAGGACAA	304
174	QY	CGAGATGGCAGTGGTTCTGAGAGAGTGTGCTACACTGTCATTATACCA---TCCCCCA	230
305	Db	TGTTGACCAGACCTACTCAGAGAGAGGTGTGCTATACCCCTCATCAATCATCGGTTCTCTG	364
231	QY	TCAGAGATCCTCCCTGAGCTCCAAATCATATGGCTATGAGAACATTGACTCCCTCACAAG	290
365	Db	TACAAGGCCATCAGGGAACTCTGCTGAAGAGTACTATGAGATGTTTCCCTCGAAGAGCTGA	424
291	QY	GAAAGTGACAGAGTTTAGAGAAAGGTTCAGAGACAGAAATATGCCCTTCT--TAGGACTTC	347

Db	425	GAGACCCAGAGAGTCTTGGAGGAAC	TGAGACTGAGTATTCACTTCTACATATGCCTTC	484
Qy	348	TGTTAGTAGCCTTGTTCCGTGCACCCATGAGCATGATTATGAAGTGTGTGTTCCACAC	405	
Db	485	TACAGACCCAGGAGTCCCGATCCCCAGAGATGATATGAATCTTCTCATGCCTCAC	542	

RESULT 11	
BM456595	
LOCUS	1070 bp
DEFINITION	linear
	EST 05-FEB-2002
	AGENCOURT 6409034 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5496261
	5', mRNA sequence.

ACCSSION	EM425232
VERSION	EM425232.1
KEYWORDS	GI:18505635
SOURCE	EST.
ORGANISM	Homo sapiens (human)
	Homo sapiens
	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE	1 (bases 1 to 1070)
AUTHORS	NIH-MGC http://mgc.nci.nih.gov/ .
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL	Unpublished (1999)
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov
Plate: LLAM12124 row: h column: 22
High quality sequence stop: 648.

```

1. .1070
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clones="IMAGE:5496251"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGC 85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH MGC Library."

```

Query Match 14.3%; Score 58; DB 12; Length 1070;
 Best Local Similarity 54.4%; Pred. No. 0.0018;
 Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2
 ORIGIN

114	QY	AAAACCTTCAGATCAAGNTAGAAAAGCCACAGAGTTTCATCCCTTCTCTAATCAGGAAA	173
306	Db	AAAGAGGCCAAGATTCCCAAAACGAAAATGAAAGAATGTCATCTCCCATCCAGGACAA	365
174	QY	CGAGAAATGGCAGTGGTCTCTGAAGAAGTGTGTACACTCTCTAATTAATCACA---TCCCCA	230
366	Db	TGTTGACCGAGACCTTACTCAGAGGAGCTGTGCTATACCCCTCATCAATCATCGGGTTCTCTG	425
221	QY	TCAGAGATCTCCCTGAGCTCCAAATGATGATGGCTATGAGAAACATTGACTCCCTCACAAG	290
426	Db	TACAAGGCCATCAGGGAACTCTGCTGAGAGTACTATGAGAATGTTCCTCGAAAGCTGA	485
291	QY	GAAGTGTAGACAGATTTTAGAGAAAGTTCAGAGACAGAATATGCCCTTCT---TAGACTTC	347
486	Db	GAGACCCAGAGATCTTGGGAGAACTGAGACTGAGTATTCACTTCTACATATGCCCTTC	545
348	QY	TGTTAGTAGGCCCTGTCTCCTCGACCCATGAGCATGATTATGAAGTTGTGTGTTCCACAC	405

Db 546 TACAGACCCAGGCGATGCCGATCCCGAAGATGAATATGAATCTTCATGCTCAC 603

RESULT 12
 BG106563 602290373F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4385178 5',
 LOCUS mRNA sequence.
 DEFINITION EST. 1095 bp mRNA linear EST 30-JAN-2001

ACCESSION
 VERSION BG106563
 KEYWORDS
 SOURCE EST. 1 GI:12600409

ORGANISM
 Homo sapiens (human)

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1095)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL0066 row: a column: 19
 High quality sequence stop: 641.
 Location/Qualifiers
 1..1095
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4385178"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

FEATURES
 source
 1..1095
 Location/Qualifiers
 1..1095
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4385178"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 14.3%; Score 58; DB 12; Length 1095;
 Best Local Similarity 54.4%; Pred. No. 0.0018;
 Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;
 QY 114 AAACTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCTATCCACTTCTAATCAGGAAA 173
 Db 207 AAAGAGGCAAGATTCCAAACGAAATGAAGAATGTCTACTCCCATCCAGGCAA 266
 QY 174 CGAGATGCGAGTGGTTCTGAAGAAGTGTGCTACACTGTCAATTAATCACA---TCCGCCA 230
 Db 267 TGTGACCAAGCTTCTAGAGGAGTGTGCTATACCTCATCATCGGGTTCTCTG 326
 QY 231 TCAGAGATCCTCCCTGAGTCCCAATGATGGCTATGAGACATTGATCCTCCCTCACAAG 290
 Db 327 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGAGCTGA 386
 QY 291 GAAGTGAAGAGTTTGAAGAGAGGTTCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
 Db 387 GAGCCAGAGAGTCTTGGAGGAAGTCTGAGACTGAGTATTCATTCTACATATGCCCTTC 446
 QY 348 TGTAGTAGGCTTGTTCCTGCAACCATGAGCATGATTAAGAGTTGTGTTTCCACAC 405
 Db 447 TACAGACCCAGGCGATGCCGATCCCGAAGATGAATATGAATCTTCATGCTCAC 504

RESULT 13
 BF795688 602259560F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342719 5',
 LOCUS mRNA sequence.
 DEFINITION EST. 1124 bp mRNA linear EST 12-JAN-2001

ACCESSION
 VERSION BF795688
 KEYWORDS
 SOURCE EST. 1 GI:12100659

ORGANISM
 Homo sapiens (human)

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1124)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL9958 row: h column: 16
 High quality sequence stop: 715.
 Location/Qualifiers
 1..1124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4342719"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

FEATURES
 source
 1..1124
 Location/Qualifiers
 1..1124
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4342719"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 14.3%; Score 58; DB 10; Length 1124;
 Best Local Similarity 54.4%; Pred. No. 0.0018;
 Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;
 QY 114 AAACTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCTATCCACTTCTAATCAGGAAA 173
 Db 282 AAAGAGGCAAGATTCCAAACGAAATGAAGAATGTCTACTCCCATCCAGGCAA 341
 QY 174 CGAGATGCGAGTGGTTCTGAAGAAGTGTGCTACACTGTCAATTAATCACA---TCCGCCA 230
 Db 342 TGTGACCAAGCTTCTAGAGGAGTGTGCTATACCTCATCATCGGGTTCTCTG 401
 QY 231 TCAGAGATCCTCCCTGAGTCCCAATGATGGCTATGAGACATTGATCCTCCCTCACAAG 290
 Db 402 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGAGCTGA 461
 QY 291 GAAGTGAAGAGTTTGAAGAGAGGTTCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
 Db 462 GAGCCAGAGAGTCTTGGAGGAAGTCTGAGACTGAGTATTCATTCTACATATGCCCTTC 521
 QY 348 TGTAGTAGGCTTGTTCCTGCAACCATGAGCATGATTAAGAGTTGTGTTTCCACAC 405
 Db 522 TACAGACCCAGGCGATGCCGATCCCGAAGATGAATATGAATCTTCATGCTCAC 579

RESULT 14
 BM455198 6405645 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500141
 LOCUS mRNA sequence.
 DEFINITION EST. 1193 bp mRNA linear EST 05-FEB-2002

ACCESSION
 VERSION BM455198
 KEYWORDS
 SOURCE BM455198

ORGANISM
 Homo sapiens (human)

REFERENCE
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
 1 (bases 1 to 1193)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLNL9958 row: h column: 16
 High quality sequence stop: 715.
 Location/Qualifiers
 1..1193
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5500141"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

FEATURES
 source
 1..1193
 Location/Qualifiers
 1..1193
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5500141"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
 Average insert size 1.867 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN
 Query Match 14.3%; Score 58; DB 10; Length 1193;
 Best Local Similarity 54.4%; Pred. No. 0.0018;
 Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;
 QY 114 AAACTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCTATCCACTTCTAATCAGGAAA 173
 Db 282 AAAGAGGCAAGATTCCAAACGAAATGAAGAATGTCTACTCCCATCCAGGCAA 341
 QY 174 CGAGATGCGAGTGGTTCTGAAGAAGTGTGCTACACTGTCAATTAATCACA---TCCGCCA 230
 Db 342 TGTGACCAAGCTTCTAGAGGAGTGTGCTATACCTCATCATCGGGTTCTCTG 401
 QY 231 TCAGAGATCCTCCCTGAGTCCCAATGATGGCTATGAGACATTGATCCTCCCTCACAAG 290
 Db 402 TACAAGGCCATCAGGGAATCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGAGCTGA 461
 QY 291 GAAGTGAAGAGTTTGAAGAGAGGTTCAGAGACAGATATGCCCTTCT---TAGGACTTC 347
 Db 462 GAGCCAGAGAGTCTTGGAGGAAGTCTGAGACTGAGTATTCATTCTACATATGCCCTTC 521
 QY 348 TGTAGTAGGCTTGTTCCTGCAACCATGAGCATGATTAAGAGTTGTGTTTCCACAC 405
 Db 522 TACAGACCCAGGCGATGCCGATCCCGAAGATGAATATGAATCTTCATGCTCAC 579

/tissue type="Peyer's patch"
/clone lib="UMNMF3"
/notes="Organ: small intestine, jejunum; Vector:
pCMVSPORT6; Site 1: NotI; Site 2: SalI; Jejunal Peyer's
patches were isolated from healthy, 4-6 week old
cross-bred pigs. RNA was extracted either immediately or
after 3 hours stimulation in an Ussing chamber with one of
the following treatments: 1) no treatment, 2) Salmonella
choleraesuis vaccine strain SC-54, 3) phorbol myristate
acetate, concanavalin A, and 8-bromo-cyclic AMP, and 4)
lipopolysaccharide and cholera toxin. Each treatment was
performed in the presence and absence of cycloheximide.
Purified poly A+ RNA from each of the treated tissues
(2-4) was combined together, reverse transcribed, and
cloned in to pCMVSPORT6 to make a library of approximately
530,000 recombinant clones with an average insert size of
1.0 kbp. Poly A+ RNA from freshly isolated and non-treated
tissue from an Ussing chamber (treatment 1) was cloned in
the same manner to produce an unstimulated cDNA library of
approximately 900,000 clones with an average insert size
of 1.5 kbp. Equal portions of the two libraries were
pooled and then subtracted with porcine ST fibroblast RNA
to create a subtracted porcine Peyer's patch library of
approximately 6,000 clones with an average insert size of
1.0 kbp."

ORIGIN

Query Match	14.2%	Score 57.4	DB 14	Length 526
Best Local Similarity	51.8%	Pred. No. 0.0024		
Matches 155	Conservative 0	Mismatches 141	Indels 3	Gaps 1

Qy	106	TTTGAAGAAACTTCACAGATCAAGATAAGAAAGCCCAAGAGTTTCATCCACTTCTAAT	165
Db	2	TTTAAAGCAACACAGATTCCTCCCAAGCAAAATGAGAAACATCATCTTCTCCCTCCAG	61
Qy	166	CAGGAAACGAGATGGCAGTGGTTCTGAGAAAGTGTCTACACTGTCATTAATCACATC	225
Db	62	GAGGAGATGTTGACCAAGCTCCGTGGAGGATCTGTCTACACCCCTCATTTGATCAGC	121
Qy	226	CCCCATCAGATCCTCCTGAGTCCCAATGATGATG---GCTATGAGAACATGACTCC	282
Db	122	ACCCCTGGAGAGGCCAATAGGCACCCCTGCTGAGGAGTGTATGAGACGTTTCCCTC	181
Qy	283	CTCACAGGAAAGTGAGACAGTTTAGAGAAAGGTCCAGAGACAGAAATATGCCCTTCTTAGG	342
Db	182	AAGACCGAGAGGCCCCAGAGAGTCATTGGGAGGACCGAGAGCTGACTCACTTATTTCAT	241
Qy	343	ACTTCTGTAGTAGCCTTGTCTGACCCATGACATGATTATGAGTTGTGTTCC	401
Db	242	GTGTCTTCTACTCTGGCCAGCCATCCTCTCAGAAATGAATGAACTTCTCATGCC	300

Search completed: April 22, 2004, 16:07:50
Job time : 2542 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 03:10:01 ; Search time 14153 Seconds
(without alignments)
11520.985 Million cell updates/sec

Title: US-10-005-907-1
Perfect score: 3762
Sequence: 1 gagaaaccagtcactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO NUC
Gapop 60.0 , Gapext 60.0

Searched: 3470272 seqs, 21671516995 residues

Word size : 0
Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : GenEmbl.*

- 1: gb.ba.*
- 2: gb.btg.*
- 3: gb.in.*
- 4: gb.cm.*
- 5: gb.ov.*
- 6: gb.pat.*
- 7: gb.ph.*
- 8: gb.pl.*
- 9: gb.pr.*
- 10: gb.ro.*
- 11: gb.sts.*
- 12: gb.sy.*
- 13: gb.un.*
- 14: gb.vi.*
- 15: em.ba.*
- 16: em.fun.*
- 17: em.hum.*
- 18: em.in.*
- 19: em.mu.*
- 20: em.om.*
- 21: em.ov.*
- 22: em.or.*
- 23: em.pat.*
- 24: em.ph.*
- 25: em.pl.*
- 26: em.ro.*
- 27: em.sts.*
- 28: em.un.*
- 29: em.vi.*
- 30: em.htg.hum.*
- 31: em.htg.inv.*
- 32: em.htg.other.*
- 33: em.htg.mus.*
- 34: em.htg.pln.*
- 35: em.htg.rod.*
- 36: em.htg.mam.*
- 37: em.htg.vrt.*
- 38: em.sv.*
- 39: em.htgo.hum.*
- 40: em.htgo_mus.*
- 41: em.htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	3762	100.0	3762	6	AX505122	AX505122 Sequence
2	3359	89.3	185467	9	AL606804	AL606804 Human DNA
3	3284	87.3	4032	9	AK126682	AK126682 Homo sapi
4	3144	83.6	4136	9	AK124520	AK124520 Homo sapi
5	3016	80.2	141268	2	AC074365	AC074365 Homo sapi
6	2506	66.6	2648	9	BC024174	BC024174 Homo sapi
7	1824	43.2	1709	9	AK123798	AK123798 Homo sapi
8	681	18.1	121353	2	AC023386	AC023386 Homo sapi
9	479	12.7	121353	2	AC023386	AC023386 Homo sapi
10	106	2.8	1924	9	AB060908	AB060908 Macaca fa
11	62	1.6	144333	9	AC105036	AC105036 Homo sapi
12	52	1.4	217	9	HS0805906	HS0805906 Homo sapi
13	52	1.4	784	5	BC052102	BC052102 Xenopus l
14	52	1.4	1815	9	BC009760	BC009760 Homo sapi
15	50	1.3	570	9	HS0807892	HS0807892 Homo sapi
16	50	1.3	1166	9	BC043545	BC043545 Homo sapi
17	50	1.3	1581	9	BC034284	BC034284 Homo sapi
18	50	1.3	1589	9	BC008899	BC008899 Homo sapi
19	50	1.3	1616	10	BC043714	BC043714 Mus muscu
20	50	1.3	1639	5	BC056324	BC056324 Danio rer
21	50	1.3	1654	9	BC040896	BC040896 Homo sapi
22	50	1.3	1737	3	AY069228	AY069228 Drosophil
23	50	1.3	1786	5	BC053205	BC053205 Danio rer
24	50	1.3	1820	9	AK026959	AK026959 Homo sapi
25	50	1.3	1880	5	BC052972	BC052972 Danio rer
26	50	1.3	2275	9	BC036869	BC036869 Homo sapi
27	50	1.3	2481	9	AB052200	AB052200 Macaca fa
28	50	1.3	3606	9	AY078404	AY078404 Homo sapi
29	50	1.3	96967	9	AC127894	AC127894 Homo sapi
30	49	1.3	159	6	AR425685	AR425685 Sequence
31	49	1.3	159	6	BD121238	BD121238 EST and e
32	49	1.3	444	9	BC038545	BC038545 Homo sapi
33	49	1.3	476	6	AX575591	AX575591 Sequence
34	49	1.3	704	9	BC063013	BC063013 Homo sapi
35	49	1.3	752	10	BC060283	BC060283 Mus muscu
36	49	1.3	959	9	S71381	S71381 prosome bet
37	49	1.3	997	9	BC032451	BC032451 Homo sapi
38	49	1.3	1008	9	BC000761	BC000761 Homo sapi
39	49	1.3	1030	9	BC029349	BC029349 Homo sapi
40	49	1.3	1051	9	BC038410	BC038410 Homo sapi
41	49	1.3	1068	9	BC004905	BC004905 Homo sapi
42	49	1.3	1109	10	BC063150	BC063150 Rattus no
43	49	1.3	1154	9	BC040307	BC040307 Homo sapi
44	49	1.3	1320	5	BC056295	BC056295 Danio rer
45	49	1.3	1346	10	BC032932	BC032932 Mus muscu
46	49	1.3	1429	6	AX587564	AX587564 Sequence
47	49	1.3	1429	3	AF097996	AF097996 Homo sapi
48	49	1.3	1440	3	AY297438	AY297438 Aiptasia
49	49	1.3	1499	9	BC042195	BC042195 Homo sapi
50	49	1.3	1595	9	BC008826	BC008826 Homo sapi
51	49	1.3	1655	9	BC039725	BC039725 Homo sapi
52	49	1.3	1661	9	BC025753	BC025753 Homo sapi
53	49	1.3	1676	6	AX188231	AX188231 Sequence
54	49	1.3	1687	10	BC035953	BC035953 Mus muscu
55	49	1.3	1706	3	AK112205	AK112205 Ciona int
56	49	1.3	1729	9	BC012854	BC012854 Homo sapi
57	49	1.3	1737	10	BC061980	BC061980 Rattus no
58	49	1.3	1786	9	AB070107	AB070107 Macaca fa
59	49	1.3	1806	3	AY118452	AY118452 Drosophil
60	49	1.3	1835	9	BC029874	BC029874 Homo sapi
61	49	1.3	1838	5	BC061320	BC061320 Silurana
62	49	1.3	1850	9	BC044938	BC044938 Homo sapi
63	49	1.3	1894	9	AB063046	AB063046 Macaca fa
64	49	1.3	1933	6	AR300635	AR300635 Sequence
65	49	1.3	1933	6	AX146886	AX146886 Sequence

301 GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGGTTCAGAGACAGAAATATGCCCTT 360
Db GACTCCCTCAAGAGAAAGTGAGACAGTTTATAGAGAAAGGTTCAGAGACAGAAATATGCCCTT 360
361 CTTAGGACTTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATATGAGTTGTG 420
Db CTTAGGACTTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATATGAGTTGTG 420
421 TTTCCACATAAATCCTCAAGCTGCTTTATCATCTTCCAGCAATGAAGCAATGAGAA 480
Db TTTCCACATAAATCCTCAAGCTGCTTTATCATCTTCCAGCAATGAAGCAATGAGAA 480
481 TAGCAGACTCTGGGAAAGTTGTTCAACCTGAGCAGTGCATGAACATCTCTTCTGGCTA 540
Db TAGCAGACTCTGGGAAAGTTGTTCAACCTGAGCAGTGCATGAACATCTCTTCTGGCTA 540
541 AAGTTTGAAGAAATATCTTATATATATATCTCTAGGCAACTCTGATATGCGCATCTCTG 600
Db AAGTTTGAAGAAATATCTTATATATATATCTCTAGGCAACTCTGATATGCGCATCTCTG 600
601 TGGCTTAGGTAATCATAGAAATGACACATGACCTTAAATATCTTATGTTGTTTTC 660
Db TGGCTTAGGTAATCATAGAAATGACACATGACCTTAAATATCTTATGTTGTTTTC 660
661 TTGTAAAGTTTGGAGACATGAGGTGATATAAAAACTTTCTTAGGCAATATATGTAATA 720
Db TTGTAAAGTTTGGAGACATGAGGTGATATAAAAACTTTCTTAGGCAATATATGTAATA 720
721 TGAATAAATTTCTAATCCCTTGACTAATGAACTGAACTCTTCTAGGCCAAAGAGA 780
Db TGAATAAATTTCTAATCCCTTGACTAATGAACTGAACTCTTCTAGGCCAAAGAGA 780
781 CCTCAGATGAACCTGAAGACTGATCTGSCCATGATAGGAAGGAGGTGAGACACACC 840
Db CCTCAGATGAACCTGAAGACTGATCTGSCCATGATAGGAAGGAGGTGAGACACACC 840
841 TTGTATACCCCTTCCCTTTTGGAGTTATGCCAAGTGACCAAGTGACCGAGTGAATAAGACT 900
Db TTGTATACCCCTTCCCTTTTGGAGTTATGCCAAGTGACCAAGTGACCGAGTGAATAAGACT 900
901 GATGAATAGACTGATTTGGCAATAGAGTCCCAATTCGAACCTGACTCTGGTGTAGAT 960
Db GATGAATAGACTGATTTGGCAATAGAGTCCCAATTCGAACCTGACTCTGGTGTAGAT 960
961 CACACACTGTCTGAGGAAATCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTT 1020
Db CACACACTGTCTGAGGAAATCCATCTATGAGACTTTGTCTACATAACAGAGACCTTGGTT 1020
1021 TCCACACCCCTTTATTTTAGCTAAGCATCTTTTCTACTGACTTCTTAAGTCTTTAGA 1080
Db TCCACACCCCTTTATTTTAGCTAAGCATCTTTTCTACTGACTTCTTAAGTCTTTAGA 1080
1081 CAAGCTTAACCTCTTTCAACCAATGCGCAATCAGACAACTTTGAATCTACTATGACCT 1140
Db CAAGCTTAACCTCTTTCAACCAATGCGCAATCAGACAACTTTGAATCTACTATGACCT 1140
1141 GTAAGCTCTCTGCTTCAAGATCTTGGCTTTTAACTGAAACCGATGCGACTTTTCCA 1200
Db GTAAGCTCTCTGCTTCAAGATCTTGGCTTTTAACTGAAACCGATGCGACTTTTCCA 1200
1201 TTTATGATTTATGCTTTGCTTGTAACTCTGCTCTCCCTAAATGTATAAAGTAAACG 1260
Db TTTATGATTTATGCTTTGCTTGTAACTCTGCTCTCCCTAAATGTATAAAGTAAACG 1260
1261 GTGACCTGACCACTCAGGCACATTTCTAGGACCTCTCTGAGAGTGTATCCAGGCGAT 1320
Db GTGACCTGACCACTCAGGCACATTTCTAGGACCTCTCTGAGAGTGTATCCAGGCGAT 1320
1321 GGTAGTCACTGCTGCTCAGATCACTCTTTAATATTTTACAGATTTGGTTTGG 1380
Db GGTAGTCACTGCTGCTCAGATCACTCTTTAATATTTTACAGATTTGGTTTGG 1380

1381 TTACCAATAGTCTCCCAAAATATATGTCGAAGAACTTCTCAATTCGAAGCTGCTCACCA 1440
Db TTACCAATAGTCTCCCAAAATATATGTCGAAGAACTTCTCAATTCGAAGCTGCTCACCA 1440
1441 AATTTCAAAATGCCAACATCTCCCATCCAAATTAATCTATTTTCACTTTGAGGTGAATCTA 1500
Db AATTTCAAAATGCCAACATCTCCCATCCAAATTAATCTATTTTCACTTTGAGGTGAATCTA 1500
1501 CTCAATAACTGTGTAGAACCCAGTGCAGACCCCTTGTCTAACCTGACATTTACTTCAAT 1560
Db CTCAATAACTGTGTAGAACCCAGTGCAGACCCCTTGTCTAACCTGACATTTACTTCAAT 1560
1561 TTTTCTTTTCTATGACTGGATATTTTTCATATATAAACTTTCGAGTAACTTCAAAAT 1620
Db TTTTCTTTTCTATGACTGGATATTTTTCATATATAAACTTTCGAGTAACTTCAAAAT 1620
1621 TAATAGTTTTGACATGCTGCTTTTCTGAGAGAGAAATGAAAGTGTCAAAAATAAAA 1680
Db TAATAGTTTTGACATGCTGCTTTTCTGAGAGAGAAATGAAAGTGTCAAAAATAAAA 1680
1681 AAGATGAATGAAGCATATATAATTTTCAATTTTCTAGTCAACAGAGATC 1740
Db AAGATGAATGAAGCATATATAATTTTCAATTTTCTAGTCAACAGAGATC 1740
1741 GAAGGATCTGTTCAAAATATTAGTAAATTAAGTAAATTAAGTAAATTAAGTAAATTTGTTT 1800
Db GAAGGATCTGTTCAAAATATTAGTAAATTAAGTAAATTAAGTAAATTTGTTT 1800
1801 GCAACACACTAGTTAATTTTAACTGTGACTAGTTATCTCCGAGGTGAGTGTAGT 1860
Db GCAACACACTAGTTAATTTTAACTGTGACTAGTTATCTCCGAGGTGAGTGTAGT 1860
1861 TTCTGTTTTTAAATTTCAAGCAAACTGGAAATAATCCATTAATTTGCTTTCTTCC 1920
Db TTCTGTTTTTAAATTTCAAGCAAACTGGAAATAATCCATTAATTTGCTTTCTTCC 1920
1921 AAGAGTTTTTAAATGATATGCGAGTCTTAAATTTGGAGCAAAAGCTTAAATGACAA 1980
Db AAGAGTTTTTAAATGATATGCGAGTCTTAAATTTGGAGCAAAAGCTTAAATGACAA 1980
1981 TGCAATTCATTATATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCCCTTAGATG 2040
Db TGCAATTCATTATATATTTTTTGTATAGTTACAGTATACGAGTTCAGTATCCCTTAGATG 2040
2041 AGATGCTGGGACAGAGAGTGTGTTGATTTGAGTTTCAAGTTATTTTGGATTTTGGATATTTT 2100
Db AGATGCTGGGACAGAGAGTGTGTTGATTTGAGTTTCAAGTTATTTTGGATTTTGGATATTTT 2100
2101 CCATACATATATGAGAGAGTTGGAAATGGAAATGGAAATGGAAATGGAAATGGAAATGGAAAT 2160
Db CCATACATATATGAGAGAGTTGGAAATGGAAATGGAAATGGAAATGGAAATGGAAATGGAAAT 2160
2161 TGTTGATATACACCTTATCTGAATAGCTGAGGTAAATTTTATACAAATTTTAAATTA 2220
Db TGTTGATATACACCTTATCTGAATAGCTGAGGTAAATTTTATACAAATTTTAAATTA 2220
2221 TTTTATGCTTGAACACAGAGTTTGGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAAT 2280
Db TTTTATGCTTGAACACAGAGTTTGGCAATTCGCAATTCGCAATTCGCAATTCGCAATTCGCAAT 2280
2281 TTCAGTCACTGCTCAAAAAGTTTTCAGATGTTTAAAGCTGATGAGTTCATGCCAGTGA 2340
Db TTCAGTCACTGCTCAAAAAGTTTTCAGATGTTTAAAGCTGATGAGTTCATGCCAGTGA 2340
2341 TCCGAGTCTTTGGGAGCCAAAGAGTGTGATCTCTTGGAGCCAGGAGTGTGAGGCCAG 2400
Db TCCGAGTCTTTGGGAGCCAAAGAGTGTGATCTCTTGGAGCCAGGAGTGTGAGGCCAG 2400
2401 ACTGCAACACAGTGTGAGACCTGTTTCTCAAAATTTTAAATAATTTAGCCAGGTGTGT 2460
Db ACTGCAACACAGTGTGAGACCTGTTTCTCAAAATTTTAAATAATTTAGCCAGGTGTGT 2460
2461 GGTGACACCTGTAGTCCCAAGGTACTCAGAGGCTGAGGTAGTAGGATTTGTTGAGACTG 2520

Db	2461	GGTGCA	CACCTGTAGTCC	CAGGTACTCAGAGG	CTCAGGTAGT	AGGATTGTTTGAGACTG	2520
Qy	2521	GGAGGTTGAGGCTGA	ACTCAGCCAGGATCTTGCCACCA	TCATCCAGCTTGGCAACAGAG	2580		
Db	2521	GGAGGTTGAGGCTGA	ACTGACCAAGATCTTGCCACCA	TCATCCAGCTTGGCAACAGAG	2580		
Qy	2581	TGAGACCCCTGTCTCA	AAAAAAGAGTTTCAGATTTTGAGAGCA	TTTCAGATCTTCAG	2640		
Db	2581	TGAGACCCCTGTCTCA	AAAAAAGAGTTTCAGATTTTGAGAGCA	TTTCAGATCTTCAG	2640		
Qy	2641	ATTAGGAGATTTTCAA	CCCTGTACTGACCTTTAGTCAATGCACAGCA	TTAATCAATAGGTG	2700		
Db	2641	ATTAGGAGATTTTCAA	CTGTACTGACCTTTAGTCAATGCACAGCA	TTAATCAATAGGTG	2700		
Qy	2701	GACTCCAGATAACTCA	ATTGCTGTATACACATTTTGCTCTCTPATTCACAGAA	TTCTTTAT	2760		
Db	2701	GACTCCAGATAACTCA	ATTGCTGTATACACATTTTGCTCTCTPATTCACAGAA	TTCTTTAT	2760		
Qy	2761	GCCTCTTGTTGGTGA	TTTAATGTCGGAGGGAACAATAGAAATTTTGCAATTC	TCTAGA	2820		
Db	2761	GCCTCTTGTTGGTGA	TTTAATGTCGGAGGGAACAATAGAAATTTTGCAATTC	TCTAGA	2820		
Qy	2821	AAAGTCATCTCTGTC	CAAAATATGTCACTCTGTAGATATTAGCCAA	TTTTAGGAAAATGAC	2880		
Db	2821	AAAGTCATCTCTGTC	CAAAATATGTCACTCTGTAGATATTAGCCAA	TTTTAGGAAAATGAC	2880		
Qy	2881	AAAATTTTTTAC	TTTTCGCTCGCTTTGACCTGTTTATGATATAAATACCTTATTTGT	2940			
Db	2881	AAAATTTTTTAC	TTTTCGCTCGCTTTGACCTGTTTATGATATAAATACCTTATTTGT	2940			
Qy	2941	AATAAAATTAATTTTA	TTTGGTGAACAACTCTGGAATTTATCAGAGAAGGGCAAGCAATA	3000			
Db	2941	AATAAAATTAATTTTA	TTTGGTGAACAACTCTGGAATTTATCAGAGAAGGGCAAGCAATA	3000			
Qy	3001	GGTTAAATAA	CAGPATTTGATGGTGAAGAAAGTTGAAATCCAGAGCA	TTCAATGCTTT	3060		
Db	3001	GGTTAAATAA	CAGPATTTGATGGTGAAGAAAGTTGAAATCCAGAGCA	TTCAATGCTTT	3060		
Qy	3061	CTGGTGCTTACCA	TATAGCCACAGCATGCTTTAATCTTTCCGAGATCTCAGTTTTTCAG	3120			
Db	3061	CTGGTGCTTACCA	TATAGCCACAGCATGCTTTAATCTTTCCGAGATCTCAGTTTTTCAG	3120			
Qy	3121	CAAAAGCAGGATTTA	AGAAATGTACTATCTTATGTGTTATGAAGAACATAGAAATCAT	3180			
Db	3121	CAAAAGCAGGATTTA	AGAAATGTACTATCTTATGTGTTATGAAGAACATAGAAATCAT	3180			
Qy	3181	GCTGTATAGTGCTTTT	TAACTGTAAATTTGTGAGCTATCTTTATGCAATAAAT	3240			
Db	3181	GCTGTATAGTGCTTTT	TAACTGTAAATTTGTGAGCTATCTTTATGCAATAAAT	3240			
Qy	3241	ATTTTGAACATTTTACA	TTTGTATATTTTAAATCAGTTTACTCAAGTGTGATATATAC	3300			
Db	3241	ATTTTGAACATTTTACA	TTTGTATATTTTAAATCAGTTTACTCAAGTGTGATATATAC	3300			
Qy	3301	AAGAAATGTAAACCA	CTGTATAGGTTATAGAAATTTGTCAAATGTATTCACCCA	3360			
Db	3301	AAGAAATGTAAACCA	CTGTATAGGTTATAGAAATTTGTCAAATGTATTCACCCA	3360			
Qy	3361	TGTAGTCACCTCTCTT	TATGAAGACAGAACACAGTACATCTCCAGAAAGTTCACAGTG	3420			
Db	3361	TGTAGTCACCTCTCTT	TATGAAGACAGAACACAGTACATCTCCAGAAAGTTCACAGTG	3420			
Qy	3421	CTCCTTTTCCCTGAG	TTTCCAGCTCTGGCAACAAATGATCTGCTTCGTATATATATAA	3480			
Db	3421	CTCCTTTTCCCTGAG	TTTCCAGCTCTGGCAACAAATGATCTGCTTCGTATATATATAA	3480			
Qy	3481	CTGTTCTAGATATTTG	TATAGCAATGTACCTTTTCCATATTTATTTTGTGTGTGAAGCTT	3540			
Db	3481	CTGTTCTAGATATTTG	TATAGCAATGTACCTTTTCCATATTTATTTTGTGTGTGAAGCTT	3540			
Qy	3541	CTTTTGTAGTCATTA	TATAATTTTGTAGATTCATCTATGTTTAAATGTTCTTATCAGTAGTTGT	3600			

3541 CTTTATAGTCATTATATAATTTTGTGAGATTCATCTATGTTTAAATGTTCTCATCAGTAGTCTT 3600

3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATACTATAATTTGTTAACTCTAATCA 3660

3601 ACATCTTACTTGTCTCAGCATATCACCATATAGATATATACTATAATTTGTTAACTCTAATCA 3660

3661 CTGATCGATATGTAGCATATTTAAAGTTTTTGCACATATCAATAAAGTGGCTATAAATGAA 3720

3661 CTGATCGATATGTAGCATATTTAAAGTTTTTGCACATATCAATAAAGTGGCTATAAATGAA 3720

3721 AA 3762

3721 AA 3762

RESULT 2

AL606804

LOCUS

DEFINITION Human DNA sequence from clone RP11-978115 on chromosome 1, complete sequence.

ACCESSION AL606804

VERSION AL606804.11 GI:24366459

KEYWORDS HTG.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS 1. (Basseg 1 to 185467)

TITLE Almeida,J.

JOURNAL Direct Submission

COMMENT Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

On Oct 24, 2002 this sequence version replaced gi:18121511.

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em: EMBL; Sw: SWISSPROT; Ir: TrEMBL; Wp: WormPEP; Information on the WormPEP database can be found at <http://www.sanger.ac.uk/Projects/C-elegans/wormpep> This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chrl>

RP11-978115 is from the library RP11-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: <http://www.sanger.ac.uk>

Contact: humquery@sanger.ac.uk

FEATURES

source

Location/Qualifiers

1. 185467

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/clones="RP11-978115"

/clone_lib="RPC1-11.4"			
ORIGIN			
Query Match 89.3%; Score 3359; DB 9; Length 185467; Best Local Similarity 99.9%; Pred. No. 0; Matches 3529; Conservative 0; Mismatches 1; Indels 1; Gaps 1;			
Qy 190	CAGGAAACGAGAAATGCGAGTGGTTCCTGAAAGAGTGTGCTACACCTGCTATTAATACATC	249	
Db 132080	CAGGAAACGAGAAATGCGAGTGGTTCCTGAAAGAGTGTGCTACACCTGCTATTAATACATC	132139	
Qy 250	CCCCATCAGAGATCCTCCCTGAGCTCCAAATGATGATGGCTATGAGACATTTGACTCCCTC	309	
Db 132140	CCCCATCAGAGATCCTCCCTGAGCTCCAAATGATGATGGCTATGAGACATTTGACTCCCTC	132199	
Qy 310	ACAAGAAAGTGAGACAGTGTATAGAGAAAGTGCAGACAGCAATATGCCCTTCTTAGGACT	369	
Db 132200	ACAAGAAAGTGAGACAGTGTATAGAGAAAGTGCAGACAGCAATATGCCCTTCTTAGGACT	132259	
Qy 370	TCTGTTAGTGGCTTCTCCCTGAGCTCCCAATGATGATGGCTATGAGCTTGGTTCCACAC	429	
Db 132260	TCTGTTAGTGGCTTCTCCCTGAGCTCCCAATGATGATGGCTATGAGCTTGGTTCCACAC	132319	
Qy 430	TAAATCCTCAAGCTGCTTTATCACCCTTCCAGCAATGAAGCAATGCAGAAATAGCAGACT	489	
Db 132320	TAAATCCTCAAGCTGCTTTATCACCCTTCCAGCAATGAAGCAATGCAGAAATAGCAGACT	132379	
Qy 490	CTGCGGAGTGTTCACCTGAGCAGTGCATGAACATTCCTTCTGGCTAAAGTTTGA	549	
Db 132380	CTGCGGAGTGTTCACCTGAGCAGTGCATGAACATTCCTTCTGGCTAAAGTTTGA	132439	
Qy 550	AATATTATCTTATATATATCCTTAGGCACTCTGATATGTCGATCTCTGTCGCTTAGG	609	
Db 132440	AATATTATCTTATATATATCCTTAGGCACTCTGATATGTCGATCTCTGTCGCTTAGG	132499	
Qy 610	TGAATCATAGAAATGACCAATGACCTAAATATTTCTATGTGTTTTCTGTTGTAAGT	669	
Db 132500	TGAATCATAGAAATGACCAATGACCTAAATATTTCTATGTGTTTTCTGTTGTAAGT	132559	
Qy 670	TTGAGGACATGCGAGTGTATGAAATGCTTTCTTAGGACCAATATGTAAGTAAATGAA	729	
Db 132560	TTGAGGACATGCGAGTGTATGAAATGCTTTCTTAGGACCAATATGTAAGTAAATGAA	132619	
Qy 730	ATTCTTAATCCCTGACTAACTGAAATGAGCCCTTCTTAGGCCAAAGACCTCAGATG	789	
Db 132620	ATTCTTAATCCCTGACTAACTGAAATGAGCCCTTCTTAGGCCAAAGACCTCAGATG	132679	
Qy 790	AACCTGAAAGACTGAATCTGCGCATGATAGGAGGGAGGTGACACACCTTGTATAC	849	
Db 132680	AACCTGAAAGACTGAATCTGCGCATGATAGGAGGGAGGTGACACACCTTGTATAC	132739	
Qy 850	CCCTTCCCTTTTGGAGTTTATGCAAGTGCACAGGATGATCATAGACTGATGAATA	909	
Db 132740	CCCTTCCCTTTTGGAGTTTATGCAAGTGCACAGGATGATCATAGACTGATGAATA	132799	
Qy 910	GACTGATTTGGCAATAGAGTCCCAATTCGAACCTGACTGCTGGTGTAGATCACAAC	969	
Db 132800	GACTGATTTGGCAATAGAGTCCCAATTCGAACCTGACTGCTGGTGTAGATCACAAC	132859	
Qy 970	TCTGAGGATTCCTATGAGACTTTGTCTACATAACAGAGACCTTGGTTTCCACAAC	1029	
Db 132860	TCTGAGGATTCCTATGAGACTTTGTCTACATAACAGAGACCTTGGTTTCCACAAC	132919	
Qy 1030	CCTTTATTTTGTAGCTTAAGCATTTCTTCTGACTTCTTAAGTCTTTAGACAAAGCTTA	1089	
Db 132920	CCTTTATTTTGTAGCTTAAGCATTTCTTCTGACTTCTTAAGTCTTTAGACAAAGCTTA	132979	
Qy 1090	ACTCTTTCAACCAATGGCAATCAGACAACTTTGAAATCTACCTATGACCTGTAGCTCT	1149	
Db 132980	ACTCTTTCAACCAATGGCAATCAGACAACTTTGAAATCTACCTATGACCTGTAGCTCT	133039	
Qy 1150	CTCCTGCTCAAGATCTTTGCTCTTTAAGCTGAACCGATGCACTTTCATTTAATGAT	1209	
Db 133040	CTCCTGCTCAAGATCTTTGCTCTTTAAGCTGAACCGATGCACTTTCATTTAATGAT	133099	
Qy 1210	TTATGCTTTTGGCTGTAACCTGCTCTCCCTAAATATGTAATAAGTAAACGGTGAACCTGA	1269	
Db 133100	TTATGCTTTTGGCTGTAACCTGCTCTCCCTAAATATGTAATAAGTAAACGGTGAACCTGA	133159	
Qy 1270	CCACCTCAGGACACCTTTCTCAGGACCTCTCAGAGTGTATCCCGAGGCATGTTAAGTCA	1329	
Db 133160	CCACCTCAGGACACCTTTCTCAGGACCTCTCAGAGTGTATCCCGAGGCATGTTAAGTCA	133219	
Qy 1330	TGTTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGGTTTACCAATA	1389	
Db 133220	TGTTGGCTCAGAAATCAACCTCTTTAAATATTTTACAGAAATTTGGGTTTGGTTTACCAATA	133279	
Qy 1390	AGTCTCCCAAAATATATGTCGAAGAACTCTCAATTCGAAGCTGCTCACCAGAAATTCAAA	1449	
Db 133280	AGTCTCCCAAAATATATGTCGAAGAACTCTCAATTCGAAGCTGCTCACCAGAAATTCAAA	133339	
Qy 1450	TGCCAAACATCTCCCCCATCAATTAATCTATTTCACTTTGAGGTGTAACTACTCAATAAA	1509	
Db 133340	TGCCAAACATCTCCCCCATCAATTAATCTATTTCACTTTGAGGTGTAACTACTCAATAAA	133399	
Qy 1510	CTGTGTAAAGACAGTGACAGACCTTTGCTGTAACCTGACATTTACTTCAATTTTCTTTT	1569	
Db 133400	CTGTGTAAAGACAGTGACAGACCTTTGCTGTAACCTGACATTTACTTCAATTTTCTTTT	133459	
Qy 1570	TCTATGTAAGTGTATTTTGCATATAAATCTGCAAGTAACTAGTTCAAAAATTAATAGTTT	1629	
Db 133460	TCTATGTAAGTGTATTTTGCATATAAATCTGCAAGTAACTAGTTCAAAAATTAATAGTTT	133519	
Qy 1630	TTGACATTTGGCTTTCTCAGAGAGAAATGAAAGTGTCAAAAATAAAAAAGATGAAA	1689	
Db 133520	TTGACATTTGGCTTTCTCAGAGAGAAATGAAAGTGTCAAAAATAAAAAAGATGAAA	133579	
Qy 1690	TGAAGCATATATAATTTGTAATTTTCAATTTTCTAGTCAACAGAGAACTCGAGGATTC	1749	
Db 133580	TGAAGCATATATAATTTGTAATTTTCAATTTTCTAGTCAACAGAGAACTCGAGGATTC	133639	
Qy 1750	TGTTCAAAATATAGTAAATAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG	1809	
Db 133640	TGTTCAAAATATAGTAAATAATGAAATGAAATGAAATGAAATGAAATGAAATGAAATG	133699	
Qy 1810	TAGTTAAATTAACCTGACTAGTTTATCTCTACCGAGGTGATGCTAGTTTCTGGTTT	1869	
Db 133700	TAGTTAAATTAACCTGACTAGTTTATCTCTACCGAGGTGATGCTAGTTTCTGGTTT	133759	
Qy 1870	TAAATTTCAAGCAAACTGGAATAATCACTCAATTAATGCTTTCTTTCCCAAGAGTTT	1929	
Db 133760	TAAATTTCAAGCAAACTGGAATAATCACTCAATTAATGCTTTCTTTCCCAAGAGTTT	133819	
Qy 1930	TTTAATGATATGCCAGCTTCTTAATTTTGGAGCAAAAGCCTTAATGCAATGCAATTCAT	1989	
Db 133820	TTTAATGATATGCCAGCTTCTTAATTTTGGAGCAAAAGCCTTAATGCAATGCAATTCAT	133879	
Qy 1990	TATATATTTTGTATAGTTTACAGTATACAGTTTGAAGTATCCCTTAGATGAGATGCTTG	2049	
Db 133880	TATATATTTTGTATAGTTTACAGTATACAGTTTGAAGTATCCCTTAGATGAGATGCTTG	133939	
Qy 2050	GGACCAAGAGTGTGGATTTTCAATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGAT	2109	
Db 133940	GGACCAAGAGTGTGGATTTTCAATTTTGGATTTTGGATTTTGGATTTTGGATTTTGGAT	133999	
Qy 2110	TAAATGAGAGTGGAAATGGATTTCAAGTCTAATCAATAAATTCACCTATGTTTGGATA	2169	
Db 134000	TAAATGAGAGTGGAAATGGATTTCAAGTCTAATCAATAAATTCACCTATGTTTGGATA	134059	
Qy 2170	TACACCTTATCTGAATAGCTGAAGTAAATTTTATCAATAATTTTAAATTAATTTATGCC	2229	
Db 134060	TACACCTTATCTGAATAGCTGAAGTAAATTTTATCAATAATTTTAAATTAATTTATGCC	134119	
Qy 2230	TGAACAGAGTTTGGCAGATTTGGACCATCAGAAAGCAGAGTGTCACTATTTCAGTCA	2289	
Db 134120	TGAACAGAGTTTGGCAGATTTGGACCATCAGAAAGCAGAGTGTCACTATTTCAGTCA	134179	

Query Match 87.3%; Score 3284; DB 9; Length 4032; Best Local Similarity 99.9%; Pred. No. 0; Matches 3604; Conservative 0; Mismatches 4; Indels 1; Gaps 1;									
QY	113	GGCAGGAATGAC	TACATTTTGAAGAAACCTTCAAGATCAAGATTAAGAAAGCCAGAG	172					
Db	405	GGCAGGAATGAC	TACATTTTGAAGAAACCTTCAAGATCAAGATTAAGAAAGCCAGAG	464					
QY	173	TTTTCATCCACTT	TAATCAGGAAACAGAGATGCGAGTGGTCTGAGAGAGTGTGTACA	232					
Db	465	TTTTCATCCACTT	TAATCAGGAAACAGAGATGCGAGTGGTCTGAGAGAGTGTGTACA	524					
QY	233	CTGTCATTAAT	CATCCCCCATCAGAGATCTCTCCCTGAGCTCCCAATGATGATGGCTATG	292					
Db	525	CTGTCATTAAT	CATCCCCCATCAGAGATCTCTCCCTGAGCTCCCAATGATGATGGCTATG	584					
QY	293	AGAACATTGACT	CCCTCACAAGGAAAGTGAGACAGATTTAGAGAAAGGTGAGAGACAGAT	352					
Db	585	AGAACATTGACT	CCCTCACAAGGAAAGTGAGACAGATTTAGAGAAAGGTGAGAGACAGAT	644					
QY	353	ATGCCCTTCTT	PAGGACTTCTGTTAGTAGGCTTGTCTCGACCCATGAGCATGATATG	412					
Db	645	ATGCCCTTCTT	PAGGACTTCTGTTAGTAGGCTTGTCTCGACCCATGAGCATGATATG	704					
QY	413	AAGTTGTGTTT	CCACATTAATCCTCAGCTGCTTTATCACCCTTCCAGCAATGAGACA	472					
Db	705	AAGTTGTGTTT	CCACATTAATCCTCAGCTGCTTTATCACCCTTCCAGCAATGAGACA	764					
QY	473	ATGAGAAATAG	CAGACTCTGCGGAGTGTTCACCCCTGAGCAGTGCATGAACATTCCTT	532					
Db	765	ATGAGAAATAG	CAGACTCTGCGGAGTGTTCACCCCTGAGCAGTGCATGAACATTCCTT	824					
QY	533	TCTGGCTAAGT	TTAGAAATATATCTTATATATCTTTAGGCACTCTGATATGTTG	592					
Db	825	TCTGGCTAAGT	TTAGAAATATATCTTATATATCTTTAGGCACTCTGATATGTTG	884					
QY	593	CATCTCTGGCT	TAGTGAATCATAGAAATGACACAAATGACCTAAAATATTTCTATGT	652					
Db	885	CATCTCTGGCT	TAGTGAATCATAGAAATGACACAAATGACCTAAAATATTTCTATGT	944					
QY	653	GTTTTTGTGTT	GAAGTTGAGGACATGGAGGTGATAAAAAACTTCTTAGGACAATA	712					
Db	945	GTTTTTGTGTT	GAAGTTGAGGACATGGAGGTGATAAAAAACTTCTTAGGACAATA	1004					
QY	713	ATGTAATAAT	GAAATTAATTTCTAATCCCTCTGACTAATCAATGAGCCCTCTCTAGGC	772					
Db	1005	ATGTAATAAT	GAAATTAATTTCTAATCCCTCTGACTAATCAATGAGCCCTCTCTAGGC	1064					
QY	773	CAAGAGACCT	CAGATGAACCTGAAAGACTGAAATTCCTGGCCATGATAGGAAGGAGTGA	832					
Db	1065	CAAGAGACCT	CAGATGAACCTGAAAGACTGAAATTCCTGGCCATGATAGGAAGGAGTGA	1124					
QY	833	GACACACCTT	GTTATACCCCTTCCCTTTGGAGTTTATGCAACAGTGACCAAGATGAGTC	892					
Db	1125	GACACACCTT	GTTATACCCCTTCCCTTTGGAGTTTATGCAACAGTGACCAAGATGAGTC	1184					
QY	893	ATAAGACTGAT	GAAATAGACTGATTTGGCAATAGAGTCCCAATTCACACTGACTCTG	952					
Db	1185	ATAAGACTGAT	GAAATAGACTGATTTGGCAATAGAGTCCCAATTCACACTGACTCTG	1244					
QY	953	GTGTAGATC	ACACTGCTGAGGATTCCTATGAGACTTTGTCTTACATACAGAGA	1012					
Db	1245	GTGTAGATC	ACACTGCTGAGGATTCCTATGAGACTTTGTCTTACATACAGAGA	1304					
QY	1013	CCTTGGTTTCC	ACACCCCTTTATTTTAGCTAAAGCATTTTCTACTGACTCTCTTAAG	1072					
Db	1305	CCTTGGTTTCC	ACACCCCTTTATTTTAGCTAAAGCATTTTCTACTGACTCTCTTAAG	1364					
QY	1073	TCCTTAGACA	AAAGCTTAATCTTTCAACCAATGGCAATCAGACAACTTTGAATCTACC	1132					
Db	1365	TCCTTAGACA	AAAGCTTAATCTTTCAACCAATGGCAATCAGACAACTTTGAATCTACC	1424					
QY	1133	TATGACCTGTA	AGCTCTCTCTGCTTCAAGATCTTTGACCTCTTTAGCTGAACCGATGTC	1192					
Db	1425	TATGACCTGTA	AGCTCTCTCTGCTTCAAGATCTTTGACCTCTTTAGCTGAACCGATGTC	1484					
QY	1193	ACTTTCCATTT	TAATGATTTATGCTTTGTTGAATCTCTCTCCCTTAAATGTATAAA	1252					
Db	1485	ACTTTCCATTT	TAATGATTTATGCTTTGTTGAATCTCTCTCCCTTAAATGTATAAA	1544					
QY	1253	AGTAAACCGT	GACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCC	1312					
Db	1545	AGTAAACCGT	GACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCC	1604					
QY	1313	CAGGCCATGT	TAAGTCATGTTGGCTCAGAACTCAACCTCTTTAAATATTTTACAGAAATTG	1372					
Db	1605	CAGGCCATGT	TAAGTCATGTTGGCTCAGAACTCAACCTCTTTAAATATTTTACAGAAATTG	1664					
QY	1373	GGTTTGGT	TACCAATAAGTCTCCAAATATATGTCACAGAACTCTCAATCCAGGCT	1432					
Db	1665	GGTTTGGT	TACCAATAAGTCTCCAAATATATGTCACAGAACTCTCAATCCAGGCT	1724					
QY	1433	GCTCACC	AAATTTCAATGCCNAATCTCCCACTCCCAATTAACCTATTTTCATCTTGAGGT	1492					
Db	1725	GCTCACC	AAATTTCAATGCCNAATCTCCCACTCCCAATTAACCTATTTTCATCTTGAGGT	1784					
QY	1493	GTAATCTACT	CAATAAACTGTGTAAAGACAGTGACAGACCCCTTTGCTAAACCTGACATTT	1552					
Db	1785	GTAATCTACT	CAATAAACTGTGTAAAGACAGTGACAGACCCCTTTGCTAAACCTGACATTT	1844					
QY	1553	ACTTCAATTT	TTCTTTCTATGTAAGTATTTTTCATATAAATTTGAGTAACTGAGTAACT	1612					
Db	1845	ACTTCAATTT	TTCTTTCTATGTAAGTATTTTTCATATAAATTTGAGTAACTGAGTAACT	1904					
QY	1613	TCAAAAATTA	ATAGTATTTTGACATTTGGCTTTCTCAGAGAGAGAAATGAAAGTGTACAA	1672					
Db	1905	TCAAAAATTA	ATAGTATTTTGACATTTGGCTTTCTCAGAGAGAGAAATGAAAGTGTACAA	1964					
QY	1673	ATAAAAAAAG	ATGAATGATATATAATTTGCAATTTTTCATATTTTCTAGTCAAC	1732					
Db	1965	ATAAAAAAAG	ATGAATGATATATAATTTGCAATTTTTCATATTTTCTAGTCAAC	2024					
QY	1733	AGAGATCGA	AGATTTCTGTTCAAAATATTAGTAAATAATGAAAAATAAACTTGTGCTTATA	1792					
Db	2025	AGAGATCGA	AGATTTCTGTTCAAAATATTAGTAAATAATGAAAAATAAACTTGTGCTTATA	2084					
QY	1793	TTTTTGTTC	CAACACTAGTTAAATTTAACCTGTGACTAGTTATCTCTACCGAAGGTGA	1852					
Db	2085	TTTTTGTTC	CAACACTAGTTAAATTTAACCTGTGACTAGTTATCTCTACCGAAGGTGA	2144					
QY	1853	TGTGTAGTTT	CTGTTTAAATTTCAAGCAAACTCGAAAAATAATCCATCTAATTTATGCTT	1912					
Db	2145	TGTGTAGTTT	CTGTTTAAATTTCAAGCAAACTCGAAAAATAATCCATCTAATTTATGCTT	2204					
QY	1913	TCCTTCCC	CAAGAGTTTAAATGATATGCCAGCTTCTCTAAATTTGGAGACAAAAGCTTGA	1972					
Db	2205	TCCTTCCC	CAAGAGTTTAAATGATATGCCAGCTTCTCTAAATTTGGAGACAAAAGCTTGA	2264					
QY	1973	ATTGCAAT	GTCAATCATATATATTTTGTATAGTTACAGTACAGTGTGAGTATCC	2032					
Db	2265	ATTGCAAT	GTCAATCATATATATTTTGTATAGTTACAGTACAGTGTGAGTATCC	2324					
QY	2033	CTTAGATG	AGATGCTGGGACAGAGTCTTTTGGATTTTTCAGATTTATTTTGGATTTTG	2092					
Db	2325	CTTAGATG	AGATGCTGGGACAGAGTCTTTTGGATTTTTCAGATTTATTTTGGATTTTG	2384					
QY	2093	GAATATTT	CCATACATATATAGAGAGTGTGAAATGGGATTCAGTCTAATCATATAA	2152					
Db	2385	GAATATTT	CCATACATATATAGAGAGTGTGAAATGGGATTCAGTCTAATCATATAA	2444					
QY	2153	TTCACTTAT	TTGATATACACTTATCTGAATAGCTCAAGGTAAATTTTATACATATTT	2212					
Db	2445	TTCACTTAT	TTGATATACACTTATCTGAATAGCTCAAGGTAAATTTTATACATATTT	2504					
QY	2213	TTAAATTA	TTTTATGCTGCAACAGAGTTTGGCACAATTTGGACCATCAGAAAGCAGAGT	2272					

Db	2446	CTTAGATGAGATGCTTGGGACCAGAGTGTGTTGGATTTTCAGATTTTATTTTGGATTTTG	2505
Qy	2093	GAATATTTCCATACATATAATAGAGAGGTTGGAAATGGGATTCAGCTCAATCATATAAA	2152
Db	2506	GAATATTTCCATACATATAATAGAGAGGTTGGAAATGGGATTCAGCTCAATCATATAAA	2565
Qy	2153	TTCACTTATGTTTGAATACACCTTATCTGAATAGCCTGGAAGGTAATTTTATACAAATTT	2212
Db	2566	TTCACTTATGTTTGAATACACCTTATCTGAATAGCCTGGAAGGTAATTTTATACAAATTT	2625
Qy	2213	TTAAATAATTTTATGCCCTGAACACAGAGTTTGGGACATTTGGACCATCAGAAAGCAGAAT	2272
Db	2626	TTAAATAATTTTATGCCCTGAACACAGAGTTTGGGACATTTGGACCATCAGAAAGCAGAAT	2685
Qy	2273	GTCACTATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGATGTCAGTTTCA	2332
Db	2686	GTCACTATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGATGTCAGTTTCA	2745
Qy	2333	GGCAGTGATCCGAGTACTTTTGGGAAGCCAGACAGGTTGGATCTCTTGGAGCCAGAGTTT	2392
Db	2746	GGCAGTGATCCGAGTACTTTTGGGAAGCCAGACAGGTTGGATCTCTTGGAGCCAGAGTTT	2805
Qy	2393	GAGGCCAGAGTGCACACACAGTGAGACCTCGTTTCTACAAATAATTAARAAATTTAGCCA	2452
Db	2806	GAGGCCAGAGTGCACACACAGTGAGACCTCGTTTCTACAAATAATTAARAAATTTAGCCA	2865
Qy	2453	GGTGTGGTGGTGCACACCTGTAGTCCAGGTAAGTCTCAGAGGCTCAGGTAGGAGTTGTT	2512
Db	2866	GGTGTGGTGGTGCACACCTGTAGTCCAGGTAAGTCTCAGAGGCTCAGGTAGGAGTTGTT	2925
Qy	2513	TTAGACTGGGAGTTGAGGCTGAACCTGAGCCAGAGATCTTGCACACATTTCCAGCTTGGG	2572
Db	2926	TTAGACTGGGAGTTGAGGCTGAACCTGAGCCAGAGATCTTGCACACATTTCCAGCTTGGG	2985
Qy	2573	CAACAGAGTGAGACCTGTCTC-----AAAAAAGTTTTCAGATTTTGGAGCAT	2627
Db	2986	CAACAGAGTGAGACCTGTCTC-----AAAAAAGTTTTCAGATTTTGGAGCAT	3045
Qy	2628	TTCAATCTTCAGATAGGATTTTCAACCTGTACTGACCTTTTAGTCATTTGCAACGAT	2687
Db	3046	TTCAATCTTCAGATAGGATTTTCAACCTGTACTGACCTTTTAGTCATTTGCAACGAT	3105
Qy	2688	TAATCAATAGGTGGACTCCAGATACTCATTTGCTGTATACACATTTTGCCTCTCTATTTC	2747
Db	3106	TAATCAATAGGTGGACTCCAGATACTCATTTGCTGTATACACATTTTGCCTCTCTATTTC	3165
Qy	2748	AACGAATTTCTATGCCCTCTTGTGGTGATTTTAAATGTCGGAAGGGAACAAATAGAAAT	2807
Db	3166	AACGAATTTCTATGCCCTCTTGTGGTGATTTTAAATGTCGGAAGGGAACAAATAGAAAT	3225
Qy	2808	TTGCAATTTCTAGAAAGTCATTTCTGCAAAATATGTCAGTCTCTGATATTTAGCCCAAT	2867
Db	3226	TTGCAATTTCTAGAAAGTCATTTCTGCAAAATATGTCAGTCTCTGATATTTAGCCCAAT	3285
Qy	2868	TTAGGAAATGACAAAATTTTACTTTTCTGCTGCTTGTAGCTTTTGTAGCTTTTATGATATA	2927
Db	3286	TTAGGAAATGACAAAATTTTACTTTTCTGCTGCTTGTAGCTTTTGTAGCTTTTATGATATA	3345
Qy	2928	ATACCTTATTTGTATATAAATTTTAAATTTGATTAACCAATCTGGAATTTATCAGAGAA	2987
Db	3346	ATACCTTATTTGTATATAAATTTTAAATTTGATTAACCAATCTGGAATTTATCAGAGAA	3405
Qy	2988	GGGCAAGCAATAGGTTTAAATCAAGTATTTGATTTGGTGAAGGACGTTGAAATCCCAAGA	3047
Db	3406	GGGCAAGCAATAGGTTTAAATCAAGTATTTGATTTGGTGAAGGACGTTGAAATCCCAAGA	3465
Qy	3048	GCATCAATGCTTCTGGTGGTTTCCACCAATGACCAAGGATGCTTAATCTTTCCGAGA	3107
Db	3466	GCATCAATGCTTCTGGTGGTTTCCACCAATGACCAAGGATGCTTAATCTTTCCGAGA	3525
Qy	3108	TCTAGTTTTTTCAGCAAGCAGGATTTAAAGATGTAATCTTATGTTGTTTATGAAGAA	3167

Db	3526	TCTAGTTTTTTCAGCAAGCAGGATTTAAGAAATGTAACATCTTATGTTGTTATGAAGAA	3585
Qy	3168	CAATAGATCATTGCTGTATAGTCTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTT	3227
Db	3586	CAATAGATCATTGCTGTATAGTCTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTT	3645
Qy	3228	TATGCATATAAATATTTGAACATTTTACATTTGTTTATATTTTAACTAGTTTACTCAAG	3287
Db	3646	TATGCATATAAATATTTGAACATTTTACATTTGTTTATATTTTAACTAGTTTACTCAAG	3705
Qy	3288	TGTGATTATATACAGAAATGTAACCACTGTAAAGGAGAGACAGACAGTACATCTCCCA	3347
Db	3706	TGTGATTATATACAGAAATGTAACCACTGTAAAGGAGAGACAGACAGTACATCTCCCA	3765
Qy	3348	ATGTATTTTCAACCATGCTAGTCACTCTCTTATGAAGAGACAGACAGTACATCTCCCA	3407
Db	3766	ATGTATTTTCAACCATGCTAGTCACTCTCTTATGAAGAGACAGACAGTACATCTCCCA	3825
Qy	3408	AGTTTCCACAGTCTCTTTTCCCTGAGTTTCCACAGTCTCTGCGCAACCAATGATCTGTT	3467
Db	3826	AGTTTCCACAGTCTCTTTTCCCTGAGTTTCCACAGTCTCTGCGCAACCAATGATCTGTT	3885
Qy	3468	CGTATAATTTAATCTGTTTCTAGATATTTTGTAGCAATGTACCTTTTCCATATTTTGT	3527
Db	3886	CGTATAATTTAATCTGTTTCTAGATATTTTGTAGCAATGTACCTTTTCCATATTTTGT	3945
Qy	3528	GTGTGTAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGATTCATCTATGTTTAAATGTT	3587
Db	3946	GTGTGTAAGGCTCTTTTGTAGTCAATTAATATTTTGTAGATTCATCTATGTTTAAATGTT	4005
Qy	3588	TATCAGTGTGTGTACATCTTCTGTCAGCATATCACCATATAGATATATATAATTT	3647
Db	4006	TATCAGTGTGTGTACATCTTCTGTCAGCATATCACCATATAGATATATATAATTT	4065
Qy	3648	GTTAATCTAATCACTGATGATATGATAGATATTTAAGTTTGTGACATATGAATAAAGT	3707
Db	4066	GTTAATCTAATCACTGATGATATGATAGATATTTAAGTTTGTGACATATGAATAAAGT	4125
Qy	3708	GGCTATAAATG 3718	
Db	4126	GGCTATAAATG 4136	

RESULT 5

AC074365/c
LOCUS
DEFINITION
AC074365
AC074365.5 GI:10280935
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 141268)
The sequence of Homo sapiens clone
Unpublished
2 (bases 1 to 141268)
Waterston, R.H.
Direct Submission
Submitted (29-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 23, 2000 this sequence version replaced gi:9838075.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: <http://genome.wustl.edu/gsc/index.shtml>
----- Project Information -----
Center project name: H_NH0115C04

----- Summary Statistics -----
Sequencing vector: M13; 100%
Chemistry: Dye-terminator; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 136274 bases at least Q40
Consensus quality: 138356 bases at least Q30
Consensus quality: 139182 bases at least Q20
Insert size: 142000; agarose-fp
Insert size: 140368; sum-of-contigs
Quality coverage: 4.61 in Q20 bases; agarose-fp
Quality coverage: 4.81 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31377: contig of 31377 bp in length
* 31378 31477: gap of unknown length
* 31478 31011: contig of 31534 bp in length
* 31011 63012: gap of unknown length
* 63012 64845: contig of 1734 bp in length
* 64845 64945: gap of unknown length
* 64945 67606: contig of 2661 bp in length
* 67606 67707: gap of unknown length
* 67707 77674: contig of 9968 bp in length
* 77674 77775: gap of unknown length
* 77775 89256: contig of 11482 bp in length
* 89256 89357: gap of unknown length
* 89357 100400: contig of 11044 bp in length
* 100400 100500: gap of unknown length
* 100500 111915: contig of 11415 bp in length
* 111915 12015: gap of unknown length
* 12015 125713: contig of 13698 bp in length
* 125713 125814: gap of unknown length
* 125814 141268: contig of 15455 bp in length.

FEATURES
source
1. 141268
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clones="RP11-115C4"
1. 31377
/note="assembly_name:Contig10"
31478. 63011
/note="assembly_name:Contig11"
63012. 64845
/note="assembly_name:Contig2"
64945. 67606
/note="assembly_name:Contig3"
67707. 77674
/note="assembly_name:Contig4"
77775. 89256
/note="assembly_name:Contig5"
89357. 100400
/note="assembly_name:Contig6"
100500. 111915
/note="assembly_name:Contig7"
112015. 125713
/note="assembly_name:Contig8"
125714. 141268
/note="assembly_name:Contig9"
clone_end:SP6
vector_side:left

ORIGIN
Query Match 80.2%; Score 3016; DB 2; Length 141268;

Best Local Similarity 99.8%; Pred. No. 0;				Matches 3526; Conservative 0; Mismatches 3; Indels 3; Gaps 3;			
QY	190	CAGGAAACGAGAAATGCGAGTGGTCTTGAAGAAAGTGTGCTACATGCTATTAATCAATC	249				
DB	5700	CAGGAAACGAGAAATGCGAGTGGTCTTGAAGAAAGTGTGCTACATGCTATTAATCAATC	5641				
QY	250	CCCATCAGAGATCCTCCCTGAGCTCCAAATGATGCTGCTATGAGAAATGATGCTCCCTC	309				
DB	5640	CCCATCAGAGATCCTCCCTGAGCTCCAAATGATGCTGCTATGAGAAATGATGCTCCCTC	5581				
QY	310	ACAAGGAAATGAGACAGTCTTATAGAAAGTGCAGAGACAGAAATATGCCCTTCTTAGACT	369				
DB	5580	ACAAGGAAATGAGACAGTCTTATAGAAAGTGCAGAGACAGAAATATGCCCTTCTTAGACT	5521				
QY	370	TCTGTACTAGGCTTGTCTCCGACCATGAGCATGATTAATGAAGTTGTGTTCCACAC	429				
DB	5520	TCTGTACTAGGCTTGTCTCCGACCATGAGCATGATTAATGAAGTTGTGTTCCACAC	5461				
QY	430	TAAATCCTCAAGCTGCTTTATCATCCCTTCCAGCAATGAAGCAATGAGAAATAGCAGACT	489				
DB	5460	TAAATCCTCAAGCTGCTTTATCATCCCTTCCAGCAATGAAGCAATGAGAAATAGCAGACT	5401				
QY	490	CTGGGAAATGTTTCAACCTGAGCAGTGCATGAACATCTCTTCTGGCTAAAGTTAGA	549				
DB	5400	CTGGGAAATGTTTCAACCTGAGCAGTGCATGAACATCTCTTCTGGCTAAAGTTAGA	5341				
QY	550	AATATTATCTTATATATATATATATATCTTAGGCAACTCTGATATGTCGCTTAGG	609				
DB	5340	AATATTATCTTATATATATATATATCTTAGGCAACTCTGATATGTCGCTTAGG	5281				
QY	610	TGAATCATAGAAATGACACATGACCTAAATATCTTATGCTTTTGTGCTTGAAGT	669				
DB	5280	TGAATCATAGAAATGACACATGACCTAAATATCTTATGCTTTTGTGCTTGAAGT	5221				
QY	670	TTGAGGACATGGAGTGTATAAATAAATCTTCTTAGGCAATATGTAATAATGAAATAA	729				
DB	5220	TTGAGGACATGGAGTGTATAAATAAATCTTCTTAGGCAATATGTAATAATGAAATAA	5161				
QY	730	ATTCTAATCCCCCTGACTAATGATGAGACCTCTCTTAGGCCAAGAGACCTCAGATG	789				
DB	5160	ATTCTAATCCCCCTGACTAATGATGAGACCTCTCTTAGGCCAAGAGACCTCAGATG	5101				
QY	790	AACCTGAAAGACTGAATCTGCGCATGATGAAGAGGAGGTGAGACACACCTTGTATAC	849				
DB	5100	AACCTGAAAGACTGAATCTGCGCATGATGAAGAGGAGGTGAGACACACCTTGTATAC	5041				
QY	850	CCCTTCCCTTTGGAGTTTATGACCAAGTACCAGGATGATCATAGACTGATGAAATA	909				
DB	5040	CCCTTCCCTTTGGAGTTTATGACCAAGTACCAGGATGATCATAGACTGATGAAATA	4981				
QY	910	GACTGATGTGCAATAAGAGTCCCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	969				
DB	4980	GACTGATGTGCAATAAGAGTCCCAATTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4921				
QY	970	TCTGAGGATTCATCTATGAGACTTTCTTACATAACAGAGACCTTGGTTTCCACACC	1029				
DB	4920	TCTGAGGATTCATCTATGAGACTTTCTTACATAACAGAGACCTTGGTTTCCACACC	4861				
QY	1030	CCTTTATTTTAGCTAAGACTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1089				
DB	4860	CCTTTATTTTAGCTAAGACTTTCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	4801				
QY	1090	ACTCTTTCAACCAATGCGCAATC-AGACAACTTTGAACTTACCTATGACCTGTAAGCTC	1148				
DB	4800	ACTCTTTCAACCAATGCGCAATC-AGACAACTTTGAACTTACCTATGACCTGTAAGCTC	4742				
QY	1149	TCTCTGCTTCAAGACTTGTGCTCTTTTAAAGCTGAACCGATGCTGCTTCAATTAAGA	1208				
DB	4741	TCTCTGCTTCAAGACTTGTGCTCTTTTAAAGCTGAACCGATGCTGCTTCAATTAAGA	4682				
QY	1209	TTTATGCTTTTGTGTAACCTCTCTCCCTTAAATGCTATATAAAGTAAACGGTGACCTG	1268				

Qy 3429 CCCTGAGTTTCACCAAGTCTGCTTGTATTAATAAAGTGTCTA 3488
 Db 2462 CCCTGAGTTTCACCAAGTCTGCTTGTATTAATAAAGTGTCTA 2403
 Qy 3489 GATATTTGTAGCAATGACCTTCCATATTTTGTGTGTAGGCTTCTTTAGT 3548
 Db 2402 GATATTTGTAGCAATGACCTTCCATATTTTGTGTGTAGGCTTCTTTAGT 2343
 Qy 3549 CATTATATATTTTGTAGATTCATCTATGTTTGTATTTCTATCATAGTGTGTACATCTTA 3608
 Db 2342 CATTATATATTTTGTAGATTCATCTATGTTTGTATTTCTATCATAGTGTGTACATCTTA 2283
 Qy 3609 CTTGTCTCAGCATATCACCATATAGATATATATAATTTGTATTAATCATCATGATGGA 3668
 Db 2282 CTTGTCTCAGCATATCACCATATAGATATATATAATTTGTATTAATCATCATGATGGA 2223
 Qy 3669 TATGTAGGATATTTAAGTTTGTGACATTTATGATTAAGTGTGCTATAAATGAA 3720
 Db 2222 TATGTAGGATATTTAAGTTTGTGACATTTATGATTAAGTGTGCTATAAATGAA 2171

RESULT 6
 BC024174 2648 bp mRNA linear PRI 04-OCT-2003
 LOCUS Homo sapiens LOC148823, mRNA (CDNA clone MGC:24564 IMAGE:4109064),
 DEFINITION complete cds.
 ACCESSION BC024174
 VERSION BC024174.1 GI:18848218
 KEYWORDS MGC.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haie, F.,
 Datchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Uedin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butlerfield, Y.S., Kzywinski, M.I., Skalska, U., Smalusz, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22398257
 12477932
 2 (bases 1 to 2648)
 Strausberg, R.
 Direct Submission
 TITLE Submitted (19-FEB-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgaps-r@mail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
 contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
 Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL Plate: 32 Row: B Column: 10
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21687148.

Location/Qualifiers
 1..2648
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:24564 IMAGE:4109064"
 /issue_type="Bone marrow, acute myelogenous leukemia"
 /clone_lib="NIH MGC_55"
 /lab_host="DH10B"
 /note="Vector: pDNR-LIB"
 1..2648
 /gene="LOC148823"
 /db_xref="LocusID:148823"
 48..455
 /codon_start=1
 /product="LOC148823"
 /protein_id="AAH24174.1"
 /db_xref="GI:18848218"
 /db_xref="LocusID:148823"
 /translation="MGNTLLRLKSLGNGKPKKGNPDDEKROEMTTFERKLODQD
 KKSQVSTSTNQENSGSEVCTVINPHQRSSLSNDGYNIDSLTRKVRQF
 RERSETVALLTSTVSRPCSTHEHDYEVVPPH"

ORIGIN

Query Match 66.6%; Score 2506; DB 9; Length 2648;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 2606; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GAGAAACCGAGTCACGTGTGAAAGATGGAAATATATCTCTCGAAATCATAGTGCCTG 60
 Db 24 GAGAAACCGAGTCACGTGTGAAAGATGGAAATATATCTCTCGAAATCATAGTGCCTG 83
 Qy 61 GAGAGAAATCAAAAGAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 120
 Db 84 GAGAGAAATCAAAAGAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 143
 Qy 121 ATGACTACATTTTGAAGAAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCA 180
 Db 144 ATGACTACATTTTGAAGAAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCA 203
 Qy 181 ACTTCTAATCAGAAAGAAAGAGATGCGAGTGTCTTCTGAAGAGTGTGTACACTGTCTAT 240
 Db 204 ACTTCTAATCAGAAAGAAAGAGATGCGAGTGTCTTCTGAAGAGTGTGTACACTGTCTAT 263
 Qy 241 AATCACATCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGCTATGAGAACATT 300
 Db 264 AATCACATCCCCATCAGAGATCCTCCCTGAGCTCCCAATGATGATGCTATGAGAACATT 323
 Qy 301 GATCCCTCACAAGGAAAGTGTAGAGAGTGTAGAGAGTGTAGAGAGTGTAGAGAGTGTAG 360
 Db 324 GATCCCTCACAAGGAAAGTGTAGAGAGTGTAGAGAGTGTAGAGAGTGTAGAGAGTGTAG 383
 Qy 361 CTTAGAGACTTCTGTAGTAGGCTTCTCTGACCCATGAGCATGATTAATGAAGTTGTG 420
 Db 384 CTTAGAGACTTCTGTAGTAGGCTTCTCTGACCCATGAGCATGATTAATGAAGTTGTG 443
 Qy 421 TTTCCACACTAAATCCTCAAGCTGTTTTATCACCTTCAGCAATGAAGCAATGAGAA 480
 Db 444 TTTCCACACTAAATCCTCAAGCTGTTTTATCACCTTCAGCAATGAAGCAATGAGAA 503
 Qy 481 TAGCAGACTCTGGGAGTGTGTTCACCTGAGCAGTGCATGAACATTCCTTCTCGCTA 540
 Db 504 TAGCAGACTCTGGGAGTGTGTTCACCTGAGCAGTGCATGAACATTCCTTCTCGCTA 563
 Qy 541 AAGTTAGAAATATATCTTATTATATATCTTCTAGGCAACTCTGATATGCGCATCTCTG 600

Db 564 AAGTTTAGAATATATATCTTATATATATCTTATGCGAATCTCTGATATGCGCACTCTG 623
Qy 601 TGGCTTAGGTGAAATCATAGAAATGACAAATGACCTTAAATATCTTATGTTGTTTTGC 660
Db 624 TGGCTTAGGTGAAATCATATATATGACAAATGACCTTAAATATCTTATGTTGTTTTGC 683
Qy 661 TTGTAAGTTTGAAGACATGAGGTGATATAAATAAATCTTCTTAGGCAATATATGTAATA 720
Db 684 TTGTAAGTTTGAAGACATGAGGTGATATAAATAAATCTTCTTAGGCAATATATGTAATA 743
Qy 721 TGAATAATAATTTCTAATCCCTGACTTAACTGTAATGACCTCTTCTTAGGCCAAAGAGA 780
Db 744 TGAATAATAATTTCTAATCCCTGACTTAACTGTAATGACCTCTTCTTAGGCCAAAGAGA 803
Qy 781 CCTCAGATGAACCTGAAAGACTGAATCTGGCCATGATAGGAAGGAGGTGAGACACACC 840
Db 804 CCTCAGATGAACCTGAAAGACTGAATCTGGCCATGATAGGAAGGAGGTGAGACACACC 863
Qy 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGCAGGATGAGTCATAAGACT 900
Db 864 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGCAGGATGAGTCATAAGACT 923
Qy 901 GATGAATAAGACTGATGTTGGCAATGAAGTCCCAATCCAACTGACTCTGGTGTAGAT 960
Db 924 GATGAATAAGACTGATGTTGGCAATGAAGTCCCAATCCAACTGACTCTGGTGTAGAT 983
Qy 961 CACACACTGCTGAGGATCCCATCTATGAGACTTTGCTATACATAACAGAGACCTTGGTT 1020
Db 984 CACACACTGCTGAGGATCCCATCTATGAGACTTTGCTATACATAACAGAGACCTTGGTT 1043
Qy 1021 TCCACAACCCCTTTATTTAGCTAAAGCAATCTTTCTACTGACTCTTAAAGTCTTTAGA 1080
Db 1044 TCCACAACCCCTTTATTTAGCTAAAGCAATCTTTCTACTGACTCTTAAAGTCTTTAGA 1103
Qy 1081 CAAAGCTTAACCTTTCAACCAATGCGCAATCAGACAACTTTGAACTTACTTACCTAGCCT 1140
Db 1104 CAAAGCTTAACCTTTCAACCAATGCGCAATCAGACAACTTTGAACTTACTTACCTAGCCT 1163
Qy 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAACCGATGTCCACTTTCCA 1200
Db 1164 GTAAGCTCTCTCTGCTTCAAGATCTTGCCTCTTTAAGCTGAACCGATGTCCACTTTCCA 1223
Qy 1201 TTTATGATTTATGCTTTGCTTTGTAATCTGCTCCCTTAAATGTAATAAGTAAACG 1260
Db 1224 TTTATGATTTATGCTTTGCTTTGTAATCTGCTCCCTTAAATGTAATAAGTAAACG 1283
Qy 1261 GTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
Db 1284 GTGACCTGACCACTCAGGCACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1343
Qy 1321 GGTAACTCATGTTGGCTCAGATCACTCTTTAAATTTTACAGAAATTTGGGTTTTGG 1380
Db 1344 GGTAACTCATGTTGGCTCAGATCACTCTTTAAATTTTACAGAAATTTGGGTTTTGG 1403
Qy 1381 TTACCAATAAGTCTCCACAAATATATGTCCAAAGAACTTTCAATCCAAAGCCTGTCTCACC 1440
Db 1404 TTACCAATAAGTCTCCACAAATATATGTCCAAAGAACTTTCAATCCAAAGCCTGTCTCACC 1463
Qy 1441 AATTTCAATGCAACATCTCCCATCCCAATTTACCTATTTTCACTTTGAGGTGTAATCTA 1500
Db 1464 AATTTCAATGCAACATCTCCCATCCCAATTTACCTATTTTCACTTTGAGGTGTAATCTA 1523
Qy 1501 CTCATAAATACTGTGAAGACCACTGACCAAGACCTTTGTAACTCCCTGACATTTACTTTCAAT 1560
Db 1524 CTCATAAATACTGTGAAGACCACTGACCAAGACCTTTGTAACTCCCTGACATTTACTTTCAAT 1583
Qy 1561 TTTTCTTTTCTATGTAAGTATTTTGGCATATTAACCTTCAGTATATAGTTCAAAAT 1620
Db 1584 TTTTCTTTTCTATGTAAGTATTTTGGCATATTAACCTTCAGTATATAGTTCAAAAT 1643
Qy 1621 TAATAGTTTTGACATTTGGCTTTCTGAGAAGAGAAATGAAAGTGTCCAAAATAAATAA 1680

RESULT 7
AK123798
LOCUS
DEFINITION Homo sapiens cDNA FLJ41804 fis, clone NOVAR2000710.
ACCESSION AK123798
VERSION AK123798.1 GI:34529425

AK123798 1709 bp mRNA linear PRI 09-SEP-2003

Db 1644 TAATAGTTTTGACATTGGCTTTTCTGAGAGAGAAATTGAAGTGTCCACAAATAAATA 1703
Qy 1681 AAGATGAATGAAGCATATATATTTTGTCAATTTTTCATTTTCTAGTCAACAGAAATC 1740
Db 1704 AAGATGAATGAAGCATATATATTTTGTCAATTTTTCATTTTCTAGCACAAGAAATC 1763
Qy 1741 GAAGGATTTCTGTTCAAAATATAGTAAATAATGAAAATAAATCTGTGCTTATATTTTGT 1800
Db 1764 GAAGGATTTCTGTTCAAAATATAGTAAATAATGAAAATAAATCTGTGCTTATATTTTGT 1823
Qy 1801 GCAACACATAGTTAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1860
Db 1824 GCAACACATAGTTAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGATGTAGT 1883
Qy 1861 TTCTGTTTTTAAATTTCAAGCAAACTGGAATAAATCCATCTAAATATGCTTTCTTTCCC 1920
Db 1884 TTCTGTTTTTAAATTTCAAGCAAACTGGAATAAATCCATCTAAATATGCTTTCTTTCCC 1943
Qy 1921 AAGAAGTTTTTAAATGATATGCCAGCTCTTAATTTGGAGACAAAGCCTTAATTGACAA 1980
Db 1944 AAGAAGTTTTTAAATGATATGCCAGCTCTTAATTTGGAGACAAAGCCTTAATTGACAA 2003
Qy 1981 TGCATTCATATATATTTTTTTTGTATAGTTACAGTATACGAGTTCAGTATCCCTTAGATG 2040
Db 2004 TGCATTCATATATATTTTTTTTGTATAGTTACAGTATACGAGTTCAGTATCCCTTAGATG 2063
Qy 2041 AGATGCTTGGGACCAAGAGTGTTTGGATTTTCAATTTTATTTTTTGGATTTTGGATTTT 2100
Db 2064 AGATGCTTGGGACCAAGAGTGTTTGGATTTTCAATTTTATTTTTTGGATTTTGGATTTT 2123
Qy 2101 CCATACATATATGAGAGAGTTGGAAATGGGATTCAGTCTAAATCATATAAATTCACCTTA 2160
Db 2124 CCATACATATATGAGAGAGTTGGAAATGGGATTCAGTCTAAATCATATAAATTCACCTTA 2183
Qy 2161 TGTTTGATATACACCTTATCTGAATAGCCTGAAGTAAATTTTATACAATTTTAAATAA 2220
Db 2184 TGTTTGATATACACCTTATCTGAATAGCCTGAAGTAAATTTTATACAATTTTAAATAA 2243
Qy 2221 TTTTATGCTGAAACAGAGTTTGGCAGCATTTGGACCATCAGAAACAGAGTGTCACTAT 2280
Db 2244 TTTTATGCTGAAACAGAGTTTGGCAGCATTTGGACCATCAGAAACAGAGTGTCACTAT 2303
Qy 2281 TTCAAGTCAAGTGTCTCAAAAGTTTTCAGATGTTAAGCTGGTGTGATGCAAGTTTCAAGTCA 2340
Db 2304 TTCAAGTCAAGTGTCTCAAAAGTTTTCAGATGTTAAGCTGGTGTGATGCAAGTTTCAAGTCA 2363
Qy 2341 TCCGAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGGAGTTTGGAGCCAG 2400
Db 2364 TCCGAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGGAGTTTGGAGCCAG 2423
Qy 2401 ACTGCAACAACACAGTGCAGACCTCGTTTTCTACAAATAATTTAAATAATTTAGCCAGGTGTGGT 2460
Db 2424 ACTGCAACAACACAGTGCAGACCTCGTTTTCTACAAATAATTTAAATAATTTAGCCAGGTGTGGT 2483
Qy 2461 GGTGCAACCTGTAGTCCAGGTTACTCAGAGGCTGAGTGTAGTGTGTTGAGACTG 2520
Db 2484 GGTGCAACCTGTAGTCCAGGTTACTCAGAGGCTGAGTGTAGTGTGTTGAGACTG 2543
Qy 2521 GGAGTTGAGGCTGAACCTGAGCCAGGATCTTGCACACATTCACGCTTTGGGCAACAGAG 2580
Db 2544 GGAGTTGAGGCTGAACCTGAGCCAGGATCTTGCACACATTCACGCTTTGGGCAACAGAG 2603
Qy 2581 TGAGACCTCTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2608
Db 2604 TGAGACCTCTCTCAAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATA 2631

oligo capping; fis (full insert sequence).

KEYWORDS

Homo sapiens (human)

SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.

REFERENCE

AUTHORS

Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T., Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R., Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiya, H., Onogawa, S., Kaeriyama, S., Satoh, N., Matsunawa, H., Takahashi, E., Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S., Teraehima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, F., Saito, H., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahari, K., Masuo, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986) NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.

FEATURES

Location/Qualifiers

1..1709

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clones="NOVAR2000710"

/tissue_type="ovary"

/clone_lib="NOVAR2"

/dev_stage="adult"

/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 43.2%; Score 1624; DB 9; Length 1709;

Best Local Similarity 99.9%; Pred. No. 0;

Matches 1674; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GAGAAACCGAGTCACTGTGAAAAGATGGGAAATATCTCTCGGAAAACTCAGTTGGCTG 60
DB 35 GAGAAACCGAGTCACTGTGAAAAGATGGGAAATATCTCTCGGAAAACTCAGTTGGCTG 94
QY 61 GGAGAGATCAAAAGAACCCAGAAAGAACCCAGATGAGGAAAGAAAGAACCCAGGAA 120
DB 95 GGAGAGATCAAAAGAACCCAGAAAGAACCCAGATGAGGAAAGAAAGAACCCAGGAA 154
QY 121 ATGACTCATTTGAAAGAAACTTCAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCC 180
DB 155 ATGACTCATTTGAAAGAAACTTCAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCC 214
QY 181 ACTTCTAATCAGGAAACGAGATGGCAGTGTTCTGGAAGATGTGTACACTGTCAAT 240
DB 215 ACTTCTAATCAGGAAACGAGATGGCAGTGTTCTGGAAGATGTGTACACTGTCAAT 274
QY 241 ATATCATCCCCCATCAGAGATCCCTCCCTGAGTCCCAATGATGATGGTATGAGAACAT 300
DB 275 ATATCATCCCCCATCAGAGATCCCTCCCTGAGTCCCAATGATGATGGTATGAGAACAT 334
QY 301 GACTCCCTCACAAGGAAAGTGAGACAGTTTATAGAGAAAGGTGAGAGACAGATATGCCCTT 360
DB 335 GACTCCCTCACAAGGAAAGTGAGACAGTTTATAGAGAAAGGTGAGAGACAGATATGCCCTT 394

18023 18122: gap of 100 bp
18123 contig of 802 bp in length
18924: gap of 100 bp
18925 contig of 805 bp in length
19029: gap of 100 bp
19230: gap of 794 bp in length
20723: contig of 100 bp
20824 20823: gap of 100 bp
21620 21619: contig of 796 bp in length
21720 21719: gap of 100 bp
22532 22531: contig of 812 bp in length
22632 22631: gap of 100 bp
23459: contig of 828 bp in length
23559: gap of 100 bp
24373: contig of 814 bp in length
24374 24473: gap of 100 bp
25288: contig of 815 bp in length
25289 25388: gap of 100 bp
25389 26192: contig of 804 bp in length
26193 26292: gap of 100 bp
26293 27089: contig of 797 bp in length
27090 27189: gap of 100 bp
28007: contig of 818 bp in length
28008 28107: gap of 100 bp
28108 28918: contig of 811 bp in length
28919 29018: gap of 100 bp
29019 29820: contig of 802 bp in length
29821 29921: gap of 100 bp
30737: contig of 817 bp in length
30738 30837: gap of 100 bp
30838 31670: contig of 833 bp in length
31671 31770: gap of 100 bp
31771 32585: contig of 815 bp in length
32586 32685: gap of 100 bp
32686 33496: contig of 811 bp in length
33497 33596: gap of 100 bp
33597 34397: contig of 801 bp in length
34398 34497: gap of 100 bp
34498 35300: contig of 803 bp in length
35301 35400: gap of 100 bp
35401 36171: contig of 771 bp in length
36172 36271: gap of 100 bp
36272 37078: contig of 807 bp in length
37079 37178: gap of 100 bp
37179 37986: contig of 808 bp in length
37987 38086: gap of 100 bp
38087 38983: contig of 797 bp in length
38984 39792: contig of 809 bp in length
39793 39892: gap of 100 bp
39893 40710: contig of 818 bp in length
40711 40810: gap of 100 bp
40811 41615: contig of 805 bp in length
41616 41715: gap of 100 bp
41716 42516: contig of 801 bp in length
42517 42616: gap of 100 bp
42617 43435: contig of 819 bp in length
43436 43535: gap of 100 bp
43536 44453: contig of 818 bp in length
44354 44453: gap of 100 bp
44454 45265: contig of 812 bp in length
45266 45365: gap of 100 bp
45366 46156: contig of 791 bp in length
46157 46256: gap of 100 bp
46257 47051: contig of 795 bp in length
47052 47151: gap of 100 bp
47152 47953: contig of 802 bp in length
47954 48053: gap of 100 bp
48054 48963: contig of 810 bp in length
48964 49863: gap of 100 bp
49864 49888: contig of 825 bp in length
49889 50706: gap of 100 bp
50706 50806: contig of 818 bp in length
50807 50707: gap of 100 bp

Query Match 18.1%; Score 681; DB 2; Length 121353;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 681; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 961 CACACACTGTCTGAGGATTCATCTATGACGACTTTGTCTACATCAACAGAGACCTTGGTT 1020
Db 4588 CACACACTGTCTGAGGATTCATCTATGACGACTTTGTCTACATCAACAGAGACCTTGGTT 4647
QY 1021 TCCACAACCCCTTTATTTTAGCTAAAGCATTTCTTTTCTACTGACTTCTTAAGTCTTTAGA 1080
Db 4648 TCCACAACCCCTTTATTTTAGCTAAAGCATTTCTTTTCTACTGACTTCTTAAGTCTTTAGA 4707
QY 1081 CAAGCTTAACCTTTTCAACCAATTCGCAATCAGACAACTTTGATCTACTATGACTTCT 1140
Db 4708 CAAGCTTAACCTTTTCAACCAATTCGCAATCAGACAACTTTGATCTACTATGACTTCT 4767
QY 1141 GTAAGCTCTCTCTGCTTCAAGATCTTGTCTCTTTAAGCTGAACCGATGTGCACCTTTCCA 1200
Db 4768 GTAAGCTCTCTCTGCTTCAAGATCTTGTCTCTTTAAGCTGAACCGATGTGCACCTTTCCA 4827
QY 1201 TTTAATGATTTATGCTTTTGTCTTGTAACTCTCTGCTCTCTTAAATGTATAAAGTAAACG 1260
Db 4828 TTTAATGATTTATGCTTTTGTCTTGTAACTCTCTGCTCTCTTAAATGTATAAAGTAAACG 4887
QY 1261 GTGACCTGACCACTCAGGCACACTTTCTCAGGACTCTCTGAGAGTGTATCCAGGCCAT 1320
Db 4888 GTGACCTGACCACTCAGGCACACTTTCTCAGGACTCTCTGAGAGTGTATCCAGGCCAT 4947
QY 1321 GGTAACTGATTTGGCTCAGAACTCAACCTTTTAAATATTTTACAGAAATTTGGGTTTGG 1380
Db 4948 GGTAACTGATTTGGCTCAGAACTCAACCTTTTAAATATTTTACAGAAATTTGGGTTTGG 5007
QY 1381 TTACCAATGATCTCCACAAATATATGTCAGAAATCTTCAATTCAGGCTTCAACCACTTCA 1440
Db 5008 TTACCAATGATCTCCACAAATATATGTCAGAAATCTTCAATTCAGGCTTCAACCACTTCA 5067
QY 1441 AATTTCAATGCAACATCTCCCAATTCATTAATTTTACAGAAATTTGGGTTTGGTATCTA 1500
Db 5068 AATTTCAATGCAACATCTCCCAATTCATTAATTTTACAGAAATTTGGGTTTGGTATCTA 5127
QY 1501 CTCATAAATGCTGTAAGACCACTGACAGACCTTTGCTTAACCTGACATTTTCAATTTCAAT 1560
Db 5128 CTCATAAATGCTGTAAGACCACTGACAGACCTTTGCTTAACCTGACATTTTCAATTTCAAT 5187
QY 1561 TTTTCTTTTCTATGACTGATTTTGTGATATAAATTTTGCATATAAATTTGCAGTAATAGTTCAAAAT 1620
Db 5188 TTTTCTTTTCTATGACTGATTTTGTGATATAAATTTTGCATATAAATTTGCAGTAATAGTTCAAAAT 5247
QY 1621 TAATAGTTTGTGACTTGGCT 1641
Db 5248 TAATAGTTTGTGACTTGGCT 5268

RESULT 9
AC023386/c 121353 bp DNA linear HTG 24-AUG-2002
LOCUS Homo sapiens chromosome 1 clone RP11-482N10 map 1, LOW-PASS
DEFINITION SEQUENCE SAMPLING.
ACCESSION AC023386
VERSION AC023386.2 GI:7417784
KEYWORDS HTG; HTGS PHASE0.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 121353)
AUTHORS Birren,B., Linton,L., Nusbaum,C. and Lander,E.
TITLE Homo sapiens chromosome 1, clone RP11-482N10
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 121353)
AUTHORS Birren,B., Linton,L., Nusbaum,C., Lander,E., Abraham,H., Allen,N.,
Anderson,S., Baldwin,J., Barna,N., Beda,F., Boguslavsky,L.,

Boukhgalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A.,
Choepl, Y., Colangelo, M., Collins, S., Collamore, A., Cooke, P.,
Dearellano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M.,
Fenster, J., Ferreira, P., Fitzhugh, M., Forrest, C., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Landers, T., Largocque, K., Lehoczyk, J., Levine, R.,
Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M.,
McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrim, J.,
Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J., Naylor, J.,
Norman, C.H., O'Connor, T., O'Donnell, P., O'Neil, D., Olivat, T.M.,
Peterson, K., Pierre, N., Pisan, C., Pollara, V., Raymond, C.,
Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S.,
Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N.,
Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Tirrell, A.,
Travers, M., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B.,
Wu, X., Wyman, D., Ye, W.J., Young, G., Zainoun, J., Zimmer, A. and
Zody, M.

Direct Submission
Submitted (14-FEB-2000) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA

3 (bases 1 to 121353)
Anderson, S., Baldwin, J., Barna, N., Bastien, V., Beda, F.,
Boguslavsky, L., Boukhgalter, B., Brown, A., Burkett, G.,
Campopiano, A., Castle, A., Choepl, Y., Colangelo, M., Collins, S.,
Collamore, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J.S.,
Dodge, S., Domino, M., Doyle, M., Ferreira, P., Fitzhugh, M., Gage, D.,
Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L.,
Grand-Pierre, N., Grant, G., Hagos, B., Heaford, A., Horton, L.,
Howland, J.C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A.,
Klein, J., Laocque, K., Lamazares, K., Landers, T., Lehoczyk, J.,
Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N.,
McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R.,
Meldrim, J., Meneus, L., Mihova, T., Miranda, C., Mienga, V., Morrow, J.,
Murphy, T., Naylor, J., Norman, C.H., O'Connor, T., O'Donnell, P.,
O'Neil, D., Olivat, T.M., Oliver, J., Peterson, K., Pierre, N.,
Pisan, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D.,
Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B.,
Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J.,
Tesfaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J.,
Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W.J.,
Young, G., Zainoun, J., Zimmer, A. and Zody, M.

Direct Submission
Submitted (24-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Apr 5, 2000 this sequence version replaced gi:6970536.
All repeats were identified using RepeatMasker:
Smit, A.P.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIER
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L5247
Center clone name: 482_N_10

NOTE: This record contains 134 individual
sequencing reads that have not been assembled into
contigs. Runs of N are used to separate the reads
and the order in which they appear is completely
arbitrary. Low-pass sequence sampling is useful for
identifying clones that may be gene-rich and allows
overlap relationships among clones to be deduced.
However, it should not be assumed that this clone
will be sequenced to completion. In the event that
the record is updated, the accession number will
be preserved.

798: contig of 798 bp in length
799 898: gap of 100 bp

1681: contig of 783 bp in length
1781: gap of 100 bp
2593: contig of 812 bp in length
2594: gap of 100 bp
3516: contig of 823 bp in length
3517: gap of 100 bp
4436: contig of 820 bp in length
4536: gap of 100 bp
5364: contig of 828 bp in length
5464: gap of 100 bp
6260: contig of 796 bp in length
6360: gap of 100 bp
7173: contig of 813 bp in length
7273: gap of 100 bp
8067: contig of 794 bp in length
8167: gap of 100 bp
8958: contig of 791 bp in length
9058: gap of 100 bp
9667: contig of 809 bp in length
9667: gap of 100 bp
10764: contig of 797 bp in length
10864: gap of 100 bp
11687: contig of 823 bp in length
11787: gap of 100 bp
12616: contig of 829 bp in length
12716: gap of 100 bp
13336: contig of 820 bp in length
13636: gap of 100 bp
14459: contig of 823 bp in length
14559: gap of 100 bp
15356: contig of 797 bp in length
15456: gap of 100 bp
16247: contig of 791 bp in length
16347: gap of 100 bp
17139: contig of 792 bp in length
17239: gap of 100 bp
18022: contig of 783 bp in length
18122: gap of 100 bp
18924: contig of 802 bp in length
19024: gap of 100 bp
19823: contig of 805 bp in length
19923: gap of 100 bp
20723: contig of 794 bp in length
20823: gap of 100 bp
21619: contig of 796 bp in length
21719: gap of 100 bp
22531: contig of 812 bp in length
22631: gap of 100 bp
23459: contig of 828 bp in length
23559: gap of 100 bp
24373: contig of 814 bp in length
24473: gap of 100 bp
25288: contig of 815 bp in length
25388: gap of 100 bp
26132: contig of 804 bp in length
26232: gap of 100 bp
27089: contig of 797 bp in length
27189: gap of 100 bp
28007: contig of 818 bp in length
28107: gap of 100 bp
28918: contig of 811 bp in length
29018: gap of 100 bp
29820: contig of 802 bp in length
29920: gap of 100 bp
30737: contig of 817 bp in length
30837: gap of 100 bp
31670: contig of 833 bp in length
31770: gap of 100 bp
32585: contig of 815 bp in length
32685: gap of 100 bp
33496: contig of 811 bp in length
33596: gap of 100 bp
34397: contig of 801 bp in length

899
1682
1782
2594
2594
3517
3517
4437
4537
5365
5465
6261
6361
7174
7274
8068
8168
8959
9059
9668
9668
10765
10865
11688
11788
12617
12717
13337
13637
14460
14560
15357
15457
16248
16348
17140
17240
18023
18123
18925
19025
19825
19925
20725
20825
21620
21720
22532
22632
23459
23559
24374
24474
25289
25389
26133
26233
27090
27190
28008
28108
28919
29019
29821
29921
30738
30838
31671
31771
32586
32686
33497
33597

	Best Local Similarity	100.0%;	Pred. No. 1.7e-41;	Mismatches	0;	Indels	0;	Gaps	0;
QY	1010	AGACCTTGGTTTCCACACCCCTTTATTTTAGCTTAAAGCACTCTTTTCTACTGACTCTTT	1069						
Db	1134	AGACCTTGGTTTCCACACCCCTTTATTTTAGCTTAAAGCACTCTTTTCTACTGACTCTTT	1193						
QY	1070	AAGCTTTAGCAAGAGCTTAACTCTTTCACCAATTCGCATCAGA	1115						
Db	1194	AAGCTTTAGCAAGAGCTTAACTCTTTCACCAATTCGCATCAGA	1239						
RESULT 11									
AC105036		144333 bp	DNA	linear	PRI 13-MAR-2003				
LOCUS									
DEFINITION									
AC105036									
VERSION									
AC105036.14		GI:28933879							
KEYWORDS									
HTG.									
SOURCE									
ORGANISM									
Homio sapiens (human)									
Homio sapiens									
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;									
Mammalia; Eutheria; Primates; Catarrhini; Homioidae; Homo.									
REFERENCE									
AUTHORS									
1 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
2 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
REFERENCE									
AUTHORS									
1 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
2 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
REFERENCE									
AUTHORS									
1 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
2 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
REFERENCE									
AUTHORS									
1 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
2 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
REFERENCE									
AUTHORS									
1 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
2 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
REFERENCE									
AUTHORS									
1 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
2 (bases 1 to 144333)									
Homio sapiens chromosome 15, clone CTD-232K18									
Unpublished									
REFERENCE									
AUTHORS									
1 (bases 1 to 144333)									

QY	2441	AAAAATTAGCCAGGTGGTGGTGCACACCTGCTACTCCAGGTACTCCAGGAGCTGAGGT	2500
Db	80097	AAAAATTAGCCAGGTGGTGGTGCACACCTGCTACTCCAGGTACTCCAGGAGCTGAGGT	80156
QY	2501	AG 2502	
Db	80157	AG 80158	
RESULT 12			
LOCUS	HSME05906	217 bp	linear
DEFINITION	Homo sapiens mRNA; cDNA DKFZp686G1888 (from clone DKFZp686G1888).		
ACCESSION	BX537812		
VERSION	BX537812.1	GI:31873786	
KEYWORDS			
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	Bukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.		
AUTHORS	1 (bases 1 to 217)		
TITLE	Wambutt, R., Heubner, D., Mewes, H.W., Weil, B., Amid, C., Osanger, A., Fobo, G., Han, M. and Wiemann, S.		
JOURNAL	Direct Submission		
COMMENT	Submitted (17-JUN-2003) MIPS, Ingolstaedter Landstr.1, D-85764 Neuherberg, GERMANY		
FEATURES	Clone from S. Wiemann, Molecular Genome Analysis, German Cancer Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de; sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing consortium of the German Genome Project.		
source	This clone (DKFZp686G1888) is available at the RZPD in Berlin. Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059 Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further information about the clone and the sequencing project is available at http://mips.gsf.de/proj/cDNA/		
Location/Qualifiers	1. .217		
/organism="Homo sapiens"			
/mol_type="mRNA"			
/db_xref="taxon:9606"			
/chromosomes="9p21.1"			
/clone="DKFZp686G1888"			
/tissue_type="human retina"			
/clone_lib="686 (synonym: hlcc3). Vector pSport1_Sfi; host DH10B; sites SfiIA + SfiIB"			
/dev_stage="adult"			
polyA signal	147. .152		
polyA_site	147. .152		
ORIGIN	166		
Query Match	1.4%; Score 52; DB 9; Length 217;		
Best Local Similarity	100.0%; Pred. No. 2.2e-14; Indels 0; Gaps 0;		
Matches	52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		
QY	3711	TATAAATGAAA 3762	
Db	164	TATAAATGAAA 215	
RESULT 13			
LOCUS	BC052102	784 bp	linear
DEFINITION	Xenopus laevis similar to anterior gradient 2 homolog (Xenopus laevis), mRNA (cDNA clone MGC:53762 IMAGE:5537106), complete cds.		
ACCESSION	BC052102		
VERSION	BC052102.1	GI:30704994	
KEYWORDS	MGC.		
SOURCE	Xenopus laevis (African clawed frog)		
ORGANISM	Xenopus laevis		
REFERENCE	1 (bases 1 to 784)		
Query Match	1.6%; Score 62; DB 9; Length 144333;		
Best Local Similarity	100.0%; Pred. No. 1.1e-19; Indels 0; Gaps 0;		
Matches	62; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		

AUTHORS
Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
TITLE
Genetic and genomic tools for Xenopus research: The NIH Xenopus Initiative
JOURNAL
Dev. Dyn. 225 (4), 384-391 (2002)
MEDLINE
22341132
PUBMED
12454917
REFERENCE
2 (bases 1 to 784)
AUTHORS
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
22388257
MEDLINE
12477932
PUBMED
3 (bases 1 to 784)
AUTHORS
Klein, S. and Strausberg, R.
TITLE
Direct Submission
JOURNAL
Submitted (01-MAY-2003) National Institutes of Health, Xenopus Gene Collection (XGC), National Institute of Child Health and Human Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD 20892-7510, USA
REFERENCE
NIH-MGC Project
Contact: XGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Igor Dawid
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Scott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 95 Row: k Column: 3
This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.
FEATURES
Location/Qualifiers
1..784
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="MGC:53762 IMAGE:5537106"
/tissue_type="embryo, stage 31/32, Xenopus"
/clone_lib="NICHD_XGC_Emb4"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1..784
gene

CDS
/gene="MGC53762"
19..513
/codon_start=1
/product="MGC53762 protein"
/protein_id="AAH52102.1"
/db_xref="GI:30704995"
/db_xref="LocusID:379407"
/translation="MESVLSIFVLLVATSTFLAKEIKPKVSKPOTLRSGWGDNLWV QYIEGLYKAKAENKPLMLNHNDCPHSLAKKAFASHQGKLAEPILNVYDP TKNKLDQGVYFKLIIFVDFSLVVRADLPKGYSNHYTFADILHLENKKALVLL KTEL"
ORIGIN
Query Match 1.4%; Score 52; DB 5; Length 784;
Best Local Similarity 100.0%; Pred. No. 1.9e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3711 TATAATGAAAAA 1815 bp mRNA linear PRI 30-SEP-2003
Db 718 TATAATGAAAAA 1815 bp mRNA linear PRI 30-SEP-2003
BC009760 Homo sapiens chromosome 4 open reading frame 9, mRNA (CDNA clone IMAGE:3936655), partial cds.
BC009760 BC009760.2 GI:33874763
Homo sapiens (human)
Homo sapiens
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1815)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stappleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1815)
Strausberg, R.
Direct Submission
Submitted (02-JUL-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:14602504.
Contact: MGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: DCTD/DTP
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Institute for Systems Biology

http://www.systemsbio.org

Contact: amadansystemsbio.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Ketteman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAL Plate: 14 Row: j Column: 14.

FEATURES

source

1..1815
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3936655"
/tissue_type="Lung, small cell carcinoma"
/clone_id="NIH_MGC_7"
/lab_host="DH10B-R-"
/note="vector: pOTB7"

gene

<1..1815
/gene="C4orf9"
/note="synonym: RES4-25"
/db_xref="LocusID:8602"

CDS

<1..1409
/gene="C4orf9"
/codon_start=3
/product="C4orf9 protein"
/protein_id="AAH09760.2"
/db_xref="GI:33874764"
/db_xref="LocusID:8602"

/translation="QAQREDALELTLEKLDODKEIOTLLSHKTPKSNRDKKPKPD
AYDMVRLGEMKQASNRMKTEABLKAEQHLKLEARLRMLCKEDENVKXP
KMSADLLNDGKQNDVQDEQKEADSPNESNEEGDSGG
EDTESPSDSHLSNVEEENEPKQRTQPKGLISGERAGKATRDLEPYT
FAPESEELSLGSRMERQLLVRIQKNHPSLAEGNAKLEKLFGLLEYVD
LATDPDLTVIDKLVHLVLCOMFPESADAIKFLVRDAMHEMIEITKGAALP
GLDVLKLTGLPPTSDFWHPVTPALVCLSQLTKPTLSLDQVVGKLFVCCFL
EVALSQRPPELINFLLGILYIATPNKASQSLVHPFRALGNKSNLLVVSAREDA
TWOQSLSLRWSRIPDLR"

misc_feature

1..1379
/gene="C4orf9"
/note="Nop14; Region: Nop14-like family. Emg1 and Nop14
are novel proteins whose interaction is required for the
maturation of the 18S rRNA and for 40S ribosome
production"
/db_xref="CDD:pfam04147"

ORIGIN

Query Match 1.4%; Score 52; DB 9; Length 1815;
Best Local Similarity 100.0%; Pred. No. 1.8e-14;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3711 TATAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

DB 1762 TATTAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1813

RESULT 15

HSM807892

LOCUS Homo sapiens mRNA; cDNA DKFZp686M0187 (from clone DKFZp686M0187).
DEFINITION BX647746.1 GI:34366903

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)

REFERENCE

1 (bases 1 to 570)
Poustka,A., Albert,R., Moosmayer,P., Schupp,I., Wellenreuther,R.,
Wewes,H.W., Weill,B., Amid,C., Osanger,A., Fobo,G., Han,M. and
Wiemann,S.

CONSRM

The German Human cDNA Consortium

TITLE

JOURNAL

COMMENT

Direct Submission
Submitted (27-AUG-2003) MIPS, Ingolstaedter Landstr.1, D-85764
Neuherberg, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
Sequenced by DKFZ (German Cancer Research Center,
Heidelberg/Germany) within the cDNA sequencing consortium of the
German Genome Project.
This clone (DKFZp686M0187) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzd.de Further
information about the clone and the sequencing project is available
at <http://mips.gsf.de/proj/cDNA/>.

FEATURES

source

1..570
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp686M0187"
/tissue_type="human retina"
/clone_lib="686 (synonym: hicc3). Vector pSport1_Sfi; host
DH10B; sites SfiIA + SfiIB"
/dev_stage="adult"
500..505
polyA_signal
polyA_site
511

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 570;
Best Local Similarity 100.0%; Pred. No. 2e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATCAAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

DB 511 TAAATCAAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 560

RESULT 16

BC043545 1166 bp mRNA linear PRI 13-JAN-2003
LOCUS Homo sapiens, clone IMAGE:5171167, mRNA.
DEFINITION BC043545
ACCESSION BC043545.1 GI:27694213
VERSION
KEYWORDS
SOURCE
ORGANISM

Homo sapiens (human)

REFERENCE

AUTHORS

TITLE

JOURNAL

Direct Submission
Submitted (09-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: The I.M.A.G.E. Consortium (LLNL)
cDNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca

REMARK

COMMENT

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smailus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found

through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 87 Row: h Column: 16.

FEATURES

Location/Qualifiers
1..1166
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5171167"
/tissue type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 1166;
Best Local Similarity 100.0%; Pred. No. 1.9e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||
DB 1031 TAAATGAAAAA 1080

RESULT 17

BC034284 1581 bp mRNA linear PRI 31-JAN-2003
LOCUS Homo sapiens, clone IMAGE:4779154, mRNA, partial cds.
DEFINITION BC034284
VERSION BC034284.1 GI:21706848

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1581)
Strausberg,R.
Direct Submission

REFERENCE

Submitted (02-JUL-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: James Cleaver, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 52 Row: d Column: 8.

FEATURES

Location/Qualifiers
1..1581
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4779154"
/tissue type="Skin, normal"
/clone_lib="NCI CGAP_Skn3"
/lab_host="DH10B"
/note="Vector: PCMV-SPORT6.ccdB"

CDS

1..546
/codon_start=1
/product="Unknown (protein for IMAGE:4779154)"
/protein_id="AAH34284.2"

/db_xref="GI:28174969"

/translation="QARLESFETLKEHVQSDYSLIMCTLOBERVRCERLEQNLDT
ELHNEILNLKQELASMEKAYQSYERARIDQEALEACQTRISKMEIQQQQVQVQV
EGLENATNLLGKIDNILLAVMVLVSVANCVVPLMKTRNRTIFSTLFLVVFIA
FLWKWDALFSYVERFFSSPR"

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 1581;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||
DB 1467 TAAATGAAAAA 1516

RESULT 18

BC008899 1589 bp mRNA linear PRI 03-OCT-2003
LOCUS Homo sapiens hypothetical protein FLJ23306, mRNA (cdna clone
MGC:14875 IMAGE:3836942), complete cds.
DEFINITION BC008899
VERSION BC008899.2 GI:33872024

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1589)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klauser,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richardson,S.,
Worley,K.C., Hale,S.S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krawinski,M.I., Skalska,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

REFERENCE

1 (bases 1 to 1589)

AUTHORS

human and mouse cDNA sequences
Generation and initial analysis of more than 15,000 full-length
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
Strausberg,R.
Direct Submission
Submitted (29-MAY-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On Aug 19, 2003 this sequence version replaced gi:14286207.
Contact: MGC help desk
Email: cgapsb@mail.nih.gov
Tissue Procurement: DCTD/DTF
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Institute for Systems Biology
<http://www.systemsbio.org>
contact: amadan@systemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettman, Anuradha
Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found

RESULT 20	BC056324	1639 bp	nrna	linear	VRT 08-OCT-2003
LOCUS	Danio rerio cDNA clone MGC:65884 IMAGE:6800994, complete cds.				
DEFINITION	BC056324				
ACCESSION	BC056324.1	GI:33604129			
VERSION	MGC.				
KEYWORDS	Danio rerio (zebrafish)				
SOURCE	Danio rerio				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.				
REFERENCE	1 (bases 1 to 1639)				
AUTHORS	Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, J., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, W., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Schecterson, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Greenwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Scheraga, A., Schein, J.E., Jones, S.J. and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22386257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1639)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgapbs@mail.nih.gov Tissue Procurement: Dr. Chi-Bin Chien cDNA Library Preparation: Invitrogen Corp DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisec.nih.gov/ Contact: nisc.mgc@nih.gov Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.				
FEATURES	source				
source	1. .1639				
organism	"Danio rerio"				
mol type	"mRNA"				

RESULT 20	BC056324	1639 bp	nrna	linear	VRT 08-OCT-2003
LOCUS	Danio rerio cDNA clone MGC:65884 IMAGE:6800994, complete cds.				
DEFINITION	BC056324				
ACCESSION	BC056324.1	GI:33604129			
VERSION	MGC.				
KEYWORDS	Danio rerio (zebrafish)				
SOURCE	Danio rerio				
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.				
REFERENCE	1 (bases 1 to 1639)				
AUTHORS	Strausberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.P., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diachenko, J., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, W., Soares, M.B., Bonaldi, M.F., Casavant, T.L., Schecterson, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.C., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Pahey, J., Helton, E., Kettner, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Greenwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E., Scheraga, A., Schein, J.E., Jones, S.J., and Marra, M.A.				
TITLE	Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences				
JOURNAL	Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)				
MEDLINE	22386257				
PUBMED	12477932				
REFERENCE	2 (bases 1 to 1639)				
AUTHORS	Strausberg, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (11-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA				
REMARK	NIH-MGC Project URL: http://mgc.nci.nih.gov				
COMMENT	Contact: MGC help desk Email: cgapbs@mail.nih.gov Tissue Procurement: Dr. Chi-Bin Chien cDNA Library Preparation: Invitrogen Corp DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL) Sequencing Center (NISC), Gaithersburg, Maryland; Web site: http://www.nisec.nih.gov/ Contact: nisc.mgc@nih.gov Akhtar, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.				
FEATURES	source				
source	1. .1639				
organism	"Danio rerio"				
mol type	"mRNA"				

```

/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN
Query Match      1.3%; Score 50; DB 9; Length 1654;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA... 1737 bp mRNA linear INV 17-DEC-2001
      |||||
      1578 TAAATGAAAAA... 1737 bp mRNA linear INV 17-DEC-2001
      |||||

RESULT 22
AY069228 Drosophila melanogaster GH27420 full length cDNA.
DEFINITION Drosophila melanogaster (fruit fly)
ACCESSION AY069228
VERSION AY069228.1 GI:17861791
KEYWORDS FTI CDNA.
SOURCE Drosophila melanogaster
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
AUTHORS Stapleton,M., Brokstein,P., Hong,L., Agbayani,A., Carlson,J.,
Champe,M., Chavez,C., Dorsett,V., Farfan,D., Frise,E., George,R.,
Gonzalez,M., Guarin,H., Li,P., Liao,G., Miranda,A., Mungall,C.J.,
Nunoo,J., Pacleb,J., Paragas,V., Park,S., Phouanenavong,S., Wan,K.,
Yu,C., Lewis,S.E., Rubin,G.M. and Celniker,S.
Direct Submission
Submitted (10-DEC-2001) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
Location/Qualifiers
1..1737
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y: cn bw sp"
/db_xref="taxon:7227"
/map="34A10-34A10"
1..1737
/gene="CG9431"
/note="alignment with genomic scaffold AB003639. gene does
not completely overlap longest ORF"
107..1345
/gene="CG9431"
/note="Longest ORF"
/codon_start=1
/product="GH27420p"
/protein_id="AAL39373.1"
/db_xref="GI:17861792"

FEATURES
source
1..1737
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y: cn bw sp"
/db_xref="taxon:7227"
/map="34A10-34A10"
1..1737
/gene="CG9431"
/note="alignment with genomic scaffold AB003639. gene does
not completely overlap longest ORF"
107..1345
/gene="CG9431"
/note="Longest ORF"
/codon_start=1
/product="GH27420p"
/protein_id="AAL39373.1"
/db_xref="GI:17861792"

CDS
107..1345
/gene="CG9431"
/note="alignment with genomic scaffold AB003639. gene does
not completely overlap longest ORF"
107..1345
/gene="CG9431"
/note="Longest ORF"
/codon_start=1
/product="GH27420p"
/protein_id="AAL39373.1"
/db_xref="GI:17861792"

ORIGIN
Query Match      1.3%; Score 50; DB 3; Length 1737;
Best Local Similarity 100.0%; Pred. No. 1.8e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA... 1786 bp mRNA linear VRT 07-OCT-2003
      |||||
      1662 TAAATGAAAAA... 1786 bp mRNA linear VRT 07-OCT-2003
      |||||

RESULT 23
BC053205 Dario rerio cDNA clone MGC:64021 IMAGE:6793127, complete cds.
LOCUS BC053205
DEFINITION Dario rerio (zebrafish)
ACCESSION BC053205
VERSION BC053205.1 GI:31418862
KEYWORDS MGC.
SOURCE Dario rerio
ORGANISM Dario rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
1 (bases 1 to 1786)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,D., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heideh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Uesdin,T.B., Tothiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahney,J., Helton,E., Kettelman,M., Madan,A., Young,A.C., Shchlenko,Y.,
Boutard,G.G., Blakeley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smallus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 1786)
Strausberg,R.
Direct Submission
Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgi.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
cDNA Sequencing by: The J.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu

```



```

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 78 Row: h Column: 3.
Location/Qualifiers
1. - 2275
/organism="Homo sapiens"
/mol_type="rRNA"
/db_xref="taxon:9606"
/clone_id="MGC:46503 IMAGE:5228549"
/tissue_type="Pancreas, Spleen, adult pooled"
/clone_lib="NIH MGC_120"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. - 2275
/gene="LOC283130"
/db_xref="LocusID:283130"
242. .982
/codon_start=1
/product="LOC283130 protein"
/protein_id="AAH36869.1"
/db_xref="GI:22477391"
/db_xref="LocusID:283130"
/translation="MWKIVRHSLGCFKMGFPIASIAVMSVLFEGVNTLLVLTAT
TQSHERRAAPSPYMHIFLAGCTGGFLQAYCLAPDILKVLONOTEPRAQPGSPPPRY
CEQVHCATKIFPEKGGPIGARGAWLTIDCTPTGVGTFVTEGLCRGYTEEGONPSSA

```



```

STS      87156. .87296
         /standard_name="RH104171"
misc_feature 91175. .92742
         /note="imperfect tandem repeat, confirmed by restriction
         map and PCR"
misc_feature 94964. .96967
         /function="unresolved tandem repeat"
         /note="overlaps bases 1. .2004 of clone AC020656"
         /function="clone overlap"

ORIGIN
Query Match      1.3%; Score 50; DB 9; Length 96967;
Best Local Similarity 100.0%; Pred. No. 1.2e-13;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2559 CATTCCAGCTGGGCAACAGAGTGAGACCCCTGCTCAAAAAAAAAAAAAA 2608
      |||
Db 38683 CATTCCAGCTGGGCAACAGAGTGAGACCCCTGCTCAAAAAAAAAAAAAA 38732

RESULT 30
AR425685      159 bp DNA linear PAT 18-DEC-2003
LOCUS
DEFINITION Sequence 17182 from patent US 6639063.
ACCESSION AR425685
VERSION AR425685.1 GI:40180795
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 159)
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 17182 28-OCT-2003;
FEATURES
    source
        1. .159
        /organism="unknown"
        /mol_type="genomic DNA"

ORIGIN
Query Match      1.3%; Score 49; DB 6; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
Db 100 AAATGAAAAA 148

RESULT 31
BD121238      159 bp DNA linear PAT 18-SEP-2002
LOCUS
DEFINITION EST and encoded human protein.
ACCESSION BD121238
VERSION BD121238.1 GI:23216148
KEYWORDS JP 2002010789-A/13315.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE
AUTHORS Edwards, J.B.D.M., Jobert, S. and Giordano, J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 13315 15-JAN-2002;
GENSET CORP
COMMENT
    OS Homo sapiens (human)
    PN JP 2002010789-A/13315
    PD 15-JAN-2002
    PF 07-AUG-2000 JP 2002080989
    PR 05-AUG-1999 US 60/147499
    PI JEAN BAPTISTE DOMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
    GIORDANO
    PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC

```

```

C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key Location/Qualifiers
FT source 1. .159
FT /organism="Homo sapiens (human)"
FEATURES
    source
        1. .159
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"

ORIGIN
Query Match      1.3%; Score 49; DB 6; Length 159;
Best Local Similarity 100.0%; Pred. No. 7.2e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
Db 100 AAATGAAAAA 148

RESULT 32
BC038545      444 bp mRNA linear PRI 15-OCT-2002
LOCUS
DEFINITION Homo sapiens, clone IMAGE:5224224, mRNA.
ACCESSION BC038545
VERSION BC038545.1 GI:23959014
KEYWORDS
SOURCE
ORGANISM
REFERENCE 1 (bases 1 to 444)
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (09-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
EC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smalus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Teal, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IMAGE Plate: 87 Row: m Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: UCSC EST-genomic
comparison.

FEATURES
    source
        1. .444
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="IMAGE:5224224"
        /tissue_type="Pancreas, Spleen, adult pooled"
        /clone_lib="NIH_MGC_120"

```

/lab_host="DH10B"
/note="vector: pCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 444;
Best Local Similarity 100.0%; Pred. No. 6.5e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

Db 390 AATGAAAAA 438

RESULT 33

AX575591/c 476 bp DNA linear PAT 07-JAN-2003
LOCUS AX575591
DEFINITION Sequence 87 from Patent WO02068645.
ACCESSION AX575591
VERSION AX575591.1 GI:27552183

KEYWORDS Homo sapiens (human)

SOURCE

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Salceda,S., Macina,R.A., Recipon,H., Cafferkey,R., Sun,Y., Liu,C.
and Turner,L.R.

TITLE Compositions and methods relating to breast specific genes and

JOURNAL

Patent: WO 02068645-A 87 06-SEP-2002;

FEATURES

Diadexus, Inc. (US)

source

1. 476
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 476;

Best Local Similarity 100.0%; Pred. No. 6.5e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

Db 126 AATGAAAAA 78

RESULT 34

BC063013 704 bp mRNA linear PRI 11-DEC-2003
LOCUS BC063013
DEFINITION Homo sapiens cDNA clone MGC:70641 IMAGE:5224517, complete cds.
ACCESSION BC063013
VERSION BC063013.1 GI:38614446

KEYWORDS

MGC.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Strausberg,R.L., Feingold,B.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Sherman,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Itoh,Y., Finkh, S.,
Carinci,P., Prange,C., Raha,S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullah,S.J., Bosak,S.A., MCEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Rulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fatey,J., Heiton,E., Kettner,M., Madan,A., Rodriguez,S., Y.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,

Bouffard G.G., Blakesley R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Sutterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.,

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 704)

Strausberg, R.

Direct Submission

TITLE

Submitted (01-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

REMARK

COMMENT

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: misc_mgc@nhgri.nih.gov

Akhter, N., Aylee, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granite, S., Guan, J., Gupta, J., Haghighi, P.,

Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 135 Row: k Column: 2

This clone was selected for full length sequencing because it

passed the following selection criteria: matched mRNA gi: 16418454.

FEATURES

Location/Qualifiers

1. 704

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:70641 IMAGE:5224517"

/tissue_type="Pancreas, Spleen, adult pooled"

/clone_lib="NIH MGC_120"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

63. 467

/codon_start=1

/product="Unknown (protein for MGC:70641)"

/protein_id="AAH63013.1"

/db_xref="GI:38614447"

/translation="MPADLSGTWLLSSDNFEGYMLAGIDPATRIAKLLKPQKIVE
QNGDSFTHTNSRLNRYFVKVGEFDEEDRLNKRCKSLVINDRLTCTIQGKEK
KXRGWTHTEGDKLHLEMPCEQVCKQTQRA"

72. 419

/note="lipocalin: Region: Lipocalin / cytosolic fatty-acid
binding protein family. Lipocalins are transporters for
small hydrophobic molecules, such as lipids, steroid
hormones, bilins, and retinoids. Alignment subsumes both
the lipocalin and fatty acid binding protein signatures
from PROSITE. This is supported on structural and
functional grounds. Structure is an eight-stranded beta
barrel"

/db_xref="CDD:pfam00061"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 704;

Best Local Similarity 100.0%; Pred. No. 6.2e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;


```

3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|||||
650 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 698

RESULT 35
BC060283 752 bp mRNA linear ROD 21-OCT-2003
LOCUS Mus musculus cDNA clone IMAGE:30291051, partial cds.
ACCESSION BC060283
VERSION BC060283.1 GI:37805219
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 752)
Strausberg,R.I., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,J., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Shat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,P.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullen,J.S., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,D.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Faney,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalios,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J., and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 752)
Strausberg,R.
Direct Submission
Submitted (20-OCT-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. David Rowe
cDNA Library Preparation: Invitrogen Corp
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC plate: 131 Row: e Column: 17
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13385879.
Location/Qualifiers
1..752

/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:30291051"
/tissue_type="Jaw and Limb, mouse, day 18.5 to new born,
pool of mature fore and hind limb, maxilla and mandible
containing endochondral and membranous bone, formed
joints, tendon, ligaments, dermis, epidermis, muscle and
teeth with newly forming dentin and enamel"
/clone_lib="NIH MGC_136"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"

ORIGIN
Query Match 1.3%; Score 49; DB 10; Length 752;
Best Local Similarity 100.0%; Pred. No. 6.2e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|||
DB 697 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 745

RESULT 36
LOCUS S71381 959 bp mRNA linear PRI 01-OCT-1994
DEFINITION prosome beta-subunit-multicatalytic proteinase complex [human,
lymphoma, mRNA, 959 nt].
ACCESSION S71381
VERSION S71381.1 GI:551546
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 959)
Gerards,W.L., Hop,F.W., Hendriks,I.L. and Bloemendal,H.
Cloning and expression of a human pro(tea)some beta-subunit cDNA: a
homologue of the yeast PRE4-subunit essential for
peptidylglutamyl-peptide hydrolase activity
FEBS Lett. 346 (2-3), 151-155 (1994)
94283586
8013624
GenBank staff at the National Library of Medicine created this
entry [NCBI gibbsq 150354] from the original journal article.
This sequence comes from Fig. 1.
Location/Qualifiers
1..959
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
1..959
/genes="prosome beta-subunit, HSBpros26"
100..801
/genes="prosome beta-subunit, HSBpros26"
/notes="multicatalytic proteinase complex; Yeast
PRE4-subunit homolog; Method: conceptual translation with
partial peptide sequencing; This sequence comes from Fig.
1; HSBpros26"
/codon_start=1
/product="prosome beta-subunit"
/protein_id="AAB31085.1"
/db_xref="GI:551547"
/translation="MDPASALYRGPIITRTQNPMTGTVLGVKFEQGVVIAADMLGSY
GSLARFNISRIKRVNNSYMLGASGDYADFOYLKQVLGOMVIDEILGDSHSYSPRAI
HSLITRWYSRKSKONPLNTWNTVIGYADGESFLGVDMLGLVAYEAPSLATGYGAYLA
QPLLRLVLEKQPVLSQTEARDLVERCMRVLIYRDARSYNRFQTATVTEKGEVEIGFLS
TETNWDIAHMSIGFE"

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 959;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```



```

/db_xref="LocusID:28981"
/translation="MKNLEVLQRRATDEMKAIVSSDQQRKKAIREQYTKNTARQEN
LGKLRKQKQVIRESHGPNKQAKWRDLQLECKQCFKQKQSTSIQGVQEGGE
DRLIL"

ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 1030;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 1051 bp mRNA linear PRI 06-OCT-2003
||||| 3714 AAATGAAAAA... 1051 bp mRNA linear PRI 06-OCT-2003
DB 948 AAATGAAAAA... 1051 bp mRNA linear PRI 06-OCT-2003

RESULT 40
BC038410 1051 bp mRNA linear PRI 06-OCT-2003
LOCUS Homo sapiens LOC92346, mRNA (cDNA clone MGC:35143 IMAGE:5169137),
DEFINITION complete cds.
ACCESSION BC038410
VERSION BC038410.1 GI:23468237
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1051)
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalek,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
PUBMED
2 (bases 1 to 1051)
Straussberg,R.
Direct Submission
Submitted (01-OCT-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: ang@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Louieged, H.,
Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 51 Row: b Column: 5
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21321731.

```

FEATURES
Location/Qualifiers
1..1051
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:35143 IMAGE:5169137"
/tissue type="Brain, adult medulla"
/clone_lib="NIH MGC_119"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1..1051
/gene="LOC92346"
/db_xref="LocusID:92346"
197..748
/codon_start=1
/product="LOC92346"
/protein_id="AAH38410.1"
/db_xref="GI:23468238"
/db_xref="LocusID:92346"
/translation="MEKRELKASVPKFDKIPWLSEASLVNKPILVSLPRYPHTSATF
LTSSKKNMLPILFQVDPVLSKARNQCDMLRNQQLCSTQCKQKQVPTKTPDD
PKASFCNMSYRMSLHQPKFQTPBPFDHDIPTENHYRLPILGPRTAHFGLITAEY
KTLKERQSSLRKEPIGKTRQ"

ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 1051;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 1051 bp mRNA linear PRI 16-SEP-2003
||||| 3714 AAATGAAAAA... 1051 bp mRNA linear PRI 16-SEP-2003
DB 976 AAATGAAAAA... 1051 bp mRNA linear PRI 16-SEP-2003

RESULT 41
BC004905 1068 bp mRNA linear PRI 16-SEP-2003
LOCUS Homo sapiens mitochondrial ribosomal protein S2, mRNA (cDNA clone
DEFINITION IMAGE:3509565), complete cds.
ACCESSION BC004905
VERSION BC004905.2 GI:33872888
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1068)
Straussberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smailus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
JOURNAL
TITLE

```

```

MEDLINE      22388257
REFERENCE    12477932
AUTHORS      Strausberg,R.
TITLE        Direct Submission
JOURNAL      Submitted (21-MAR-2001) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       On Aug 19, 2003 this sequence version replaced gi:13436187.
              Contact: MGC help desk
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: ATCC
              cDNA Library Preparation: Rubin Laboratory
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Institute for Systems Biology
              http://www.systemsbio.org
              Contact: amadan@systemsbiology.org
              Anup Madan, Jessica Fahay, Erin Helton, Mark Kettelman, Anuradha
              Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 11 Row: 0 Column: 7
This clone has the following problem: The cds is short compared to
the longest cds in the locus.

FEATURES             source
1..1068
   /organism="Homo sapiens"
   /mol_type="mRNA"
   /db_xref="taxon:9606"
   /clone="IMAGE:3509565"
   /tissue_type="Brain, neuroblastoma"
   /clone_lib="NIH MGC 19"
   /lab_host="DH10B-R"
   /note="Vector: pOTB7"
1..1068
   /gene="MRP2"
   /note="synonyms: CGI-91, MRP-S2"
   /db_xref="LocusID:51116"
291..452
   /codon_start=1
   /product="MRP2 protein"
   /protein_id="AAH04905.2"
   /db_xref="GI:33872889"
   /db_xref="LocusID:51116"
   /translation="MTTLRLWLTCTAGSSRRSPGPRRGSRRLSLIACRRARSPGTR
GQPTLLGLT"

ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 1068;
Best Local Similarity 100.0%; Pred. No. 6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||
Db 999 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1047

RESULT 42
BC063150
LOCUS          BC063150          1109 bp      mRNA      linear      ROD 11-DEC-2003
DEFINITION    Rattus norvegicus cDNA clone MGC:72699 IMAGE:6922267, complete cds.
ACCESSION     BC063150
VERSION       BC063150.1 GI:38649091
KEYWORDS      MGC.
SOURCE        Rattus norvegicus (Norway rat)
ORGANISM      Rattus norvegicus
              Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;
              Rattus.
REFERENCE     1 (bases 1 to 1109)

```

```

AUTHORS      Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
              Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
              Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
              Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
              Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
              Scapleton,M., Soares,M.B., Bonaldo,M.P., Casavant,T.L.,
              Schretz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
              Carrinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
              Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
              McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
              Villalon,D.K., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
              Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodrigues,S.,
              Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
              Buffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
              Dutton,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
              Bicknell,Y.S., Krzywinski,M.I., Skalska,U., Smalish,D.E.,
              Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
              Generation and initial analysis of more than 15,000 full-length
              human and mouse cDNA sequences
              Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
              12477932
2 (bases 1 to 1109)
Direct Submission
Submitted (02-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgapbs-r@mail.nih.gov
              Tissue Procurement: John C. Marshall, M.D., Ph.D
              cDNA Library Preparation: CLONTECH Laboratories, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
              DNA Sequencing by: Genome Sequence Centre,
              BC Cancer Agency, Vancouver, BC, Canada
              info@cgsc.bc.ca
              Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
              Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
              Letticia Hsiao, Martin Krzywinski, Reta Kutche, Oliver Lee, Soo
              Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
              Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saedi, Jacqueline
              Schein, Duane Smalish, Michael Smith, Lorraine Spence, Jeff Scott,
              Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
              George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 52 Row: 6 Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 19705452.

FEATURES             Location/Qualifiers
1..1109
   /organism="Rattus norvegicus"
   /mol_type="mRNA"
   /db_xref="taxon:10116"
   /clone="MGC:72699 IMAGE:6922267"
   /tissue_type="Pituitary gland, anterior, rat"
   /clone_lib="NICHD Rr Pit1"
   /lab_host="DH10B"
   /note="Vector: pDNR-LIB"
70..537
   /codon_start=1
   /product="Unknown (protein for MGC:72699)"
   /protein_id="AAH63150.1"
   /db_xref="GI:38649092"
   /translation="MADLKQMDNEVLAFTSYATIILAKOMFLSSATQRLTNKVF
ANPDCAFGKGENAKFLRDEKVRRAHLNDLENIVPFLGILSLSGDLSST
ALIHFRIFVGARIYHTIAYLTFLPQPNRGLAFLVGVGYVTLMSAYLLRSRLYL"
94..531
   /note="MAPEG; Region: MAPEG family. This family is has
              been called MAPEG (Membrane Associated Proteins in

```

Eicosanoid and Glutathione metabolism). It includes proteins such as prostaglandin H synthase. This enzyme catalyses the synthesis of PGE2 from PGH2 (produced by cyclooxygenase from arachidonic acid). Because of structural similarities in the active sites of FLAP, LTC4 synthase and PGE synthase, substrates for each enzyme can compete with one another and modulate synthetic activity" /db_xref="CDD:pfam01124"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 1109;
Best Local Similarity 100.0%; Pred. No. 5.9e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGCAA 3762
DB 893 AATGCAA 941

RESULT 43

BC040907 1154 bp mRNA linear PRI 11-DEC-2002
LOCUS Homo sapiens, clone IMAGE:5744167, mRNA.
DEFINITION BC040907
ACCESSION BC040907.1 GI:26454774
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 1154)
Direct Submission
Strausberg, R.

AUTHORS

Submitted (06-DEC-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www.shgc.stanford.edu>
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 88 Row: 1 Column: 1.

FEATURES

source

1..1154
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5744167"
/tissue_type="Brain, adult medulla"
/clone_lib="NIH MGC 119"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1154;
Best Local Similarity 100.0%; Pred. No. 5.9e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGCAA 3762
DB 1090 AATGCAA 1138

RESULT 44

BC056295
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

REMARK

COMMENT

FEATURES

source

ORGANISM="Danio rerio"

BC056295 1320 bp mRNA linear VRT 08-OCT-2003
Danio rerio cDNA clone MGC:65806 IMAGE:6792135, complete cds.
BC056295
BC056295.1 GI:33604046
MGC
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes; Cyprinidae; Danio.

1 (bases 1 to 1320)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, S.I., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaby, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., R.A., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalls, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A. M.A. full-length generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2 (bases 1 to 1320)
Strausberg, R.
Direct Submission
Submitted (11-AUG-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@hgr.nih.gov
Alkhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., Mcbowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Turgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 121 Row: e Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source

ORGANISM="Danio rerio"

```

/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:65806 IMAGE:6792135"
/tissue_type="kidney, zebrafish"
/clone_lib="NCI CGAP_2K1d1"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
124.933
/codon_start=1
/product="Unknown (protein for MGC:65806)"
/protein_id="AAH56295.1"
/db_xref="GI:33604047"
/translation="MGQEDMSAEDLADQFLVTKOVLPHWRLICLISTPLEDGIWM
WQWSEQDYIATWSCGFLATCFVILNLGQGGCVLRLNLVQVACPLGICIIA
LQTVAYSILMDKFLMNLALGSLAASRSGKSFAGVFSMGSFQKQNLG
GNVLIVLMTLLHSDFFSILQNMVGTILVAVGFKTGLAALTIVWLLAINV
FNFAFTVAYKPMHDFLKYDFQTTSVIGLLLVVAGPGGVSMDKKEW"
133.930
misc_feature
133.930
/notes="SURF4; Region: SURF4 family"
/db_xref="CDD:pfam02077"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 1320;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
DB 1267 AAATGAAAAA... 1315

RESULT 45
BC032932 1346 bp mRNA linear ROD 06-OCT-2003.
LOCUS BC032932
DEFINITION Mus musculus cDNA clone MGC:41415 IMAGE:1529010, complete cds.
ACCESSION BC032932.1 GI:21410723
VERSION BC032932.1
KEYWORDS MGC.
SOURCE Mus musculus
ORGANISM Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 1346)
Klausner,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klauser,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneitz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Roha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodrigues,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.C., Blakesley,R.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
MEDLINE
PUBMED
12477932
REFERENCE
2 (bases 1 to 1346)
Strausberg,R.
Direct Submission
Submitted (10-JUN-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

```

```

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Marcello Bento Soares, Ph.D.
cDNA Library Preparation: M. Bento Soares, University of Iowa
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Liisa Prabh, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalhus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Marasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Nacso Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 67 Row: k Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: Genomescan gene
prediction, similarity but not identity to protein.

FEATURES
Location/Qualifiers
1..1346
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="MGC:41415 IMAGE:1529010"
/tissue_type="Mammary gland, lactating mouse"
/clone_lib="Soares_mammary_gland_NMLMG"
/lab_host="DH10B"
/notes="Vector: pT7T3-Pac"
135.569
/codon_start=1
/product="Unknown (protein for MGC:41415)"
/protein_id="AAH32932.1"
/db_xref="GI:21410724"
/translation="MALLVMEDEEDSKHFNVDKIVHONLSKKKKQMKKKKELVED
DFEVNSDARFQMYTSHLFLNLDPSDFNFKTKAMEKILEKARHREKBEILLIQAIVE
RAQQDTGKTQKQMPDPALESLIKSVKNTEQFOARKQKRVK"
144.380
/misc_feature
144.380
/notes="COG5638; Region: Uncharacterized conserved protein
[Function unknown]"
/db_xref="CDD:COG5638"

ORIGIN
Query Match 1.3%; Score 49; DB 10; Length 1346;
Best Local Similarity 100.0%; Pred. No. 5.8e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
DB 1291 AAATGAAAAA... 1339

RESULT 46
AX587564 1429 bp DNA linear PAT 10-JAN-2003
LOCUS AX587564
DEFINITION Sequence 34 from Patent WO0246467.
ACCESSION AX587564
VERSION AX587564.1 GI:28212327
KEYWORDS
SOURCE synthetic construct
ORGANISM synthetic construct
artificial sequences.
1
REFERENCE
1
Bertucci,F., Houlgatte,R., Birnbaum,D., Nguyen,C., Viens,P. and
Fert,V.
Gene expression profiling of primary breast carcinomas using arrays
of candidate genes
Patent: WO 0246467-A 34 13-JUN-2002;
JOURNAL

```

```

FEATURES             source
  Ipoegen (FR)
  Location/Qualifiers
    1..1429
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /note="primer"
  misc_feature
    1..1429
    /note="protein phosphatase 4 (formerly x), catalytic
    subunit (PPP4C) gene."
  ORIGIN
    Query Match      1.3%; Score 49; DB 6; Length 1429;
    Best Local Similarity 100.0%; Pred. No. 5.8e-13;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 3714 AAATGAAAAA... 1429 bp mRNA linear PRI 17-DEC-1998
  Db 1336 AAATGAAAAA... 1429 bp mRNA linear PRI 17-DEC-1998
  RESULT 47
  LOCUS AF097996 1429 bp mRNA linear PRI 17-DEC-1998
  DEFINITION Homo sapiens protein phosphatase X (PPX) mRNA, complete cds.
  ACCESSION AF097996
  VERSION AF097996.1 GI:4028574
  KEYWORDS
  SOURCE Homo sapiens (human)
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  AUTHORS Hu.M.C., Tang-Oxley.Q., Qiu.W.R., Wang.Y.P.,
  Mihindukulasuriya.K.A., Afshar.R. and Tan.T.H.
  TITLE Protein phosphatase X interacts with c-Rel and stimulates
  c-Rel/nuclear factor kappaB activity
  JOURNAL J. Biol. Chem. 273 (50), 33561-33565 (1998)
  MEDLINE 99057922
  PUBMED 9837938
  REFERENCE
  AUTHORS Hu.M.C.-T., Tang-Oxley.Q., Qiu.W., Wang.Y.-P.,
  Mihindukulasuriya.K.A., Afshar.R. and Tan.T.-H.
  TITLE Direct Submission
  JOURNAL Submitted (13-OCT-1998) Department of Cell Biology, Amgen, Inc.,
  One Amgen Center Road, 141-D, Thousand Oaks, CA 91320, USA
  FEATURES             source
    Location/Qualifiers
      1..1429
      /organism="Homo sapiens"
      /mol_type="mRNA"
      /db_xref="taxon:9606"
      1..1429
      /gene="PPX"
      105..1028
      /gene="PPX"
      /codon_start=1
      /product="protein phosphatase X"
      /protein_id="AAC96318.1"
      /db_xref="GI:4028575"
      /translation="MAESDLDRIQLRCELKESVAKALCAKARELLVESNVOR
      VDSPTVCGDILHGQSDIKELFRVGVGVFETNYLFGMDVDRGFYSVETILLALKV
      RYPRDITLLIHGHSRQITQYGFYDECLRGYFVWRYCTEIDFYLISALIDGKI
      FCVHGGLSPSIQTLDIIRDKQVHPGMDLLMSDPEDTGWGSPRGAGYLF
      SDVAQFNAANDIMICRAHQLVMEGVKWFNETVLTWSAPNYCYRCGNVAALLELD
      EHLQKDFIIFEAAPQETRGIPSKPVDYFL"
  ORIGIN
    Query Match      1.3%; Score 49; DB 9; Length 1429;
    Best Local Similarity 100.0%; Pred. No. 5.8e-13;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 3714 AAATGAAAAA... 1429 bp mRNA linear PRI 17-DEC-1998

```

```

Db 1336 AAATGAAAAA... 1440 bp mRNA linear INV 01-NOV-2003
  RESULT 48
  LOCUS AY297438 1440 bp mRNA linear INV 01-NOV-2003
  DEFINITION Aiptasia pulchella beta actin mRNA, complete cds.
  ACCESSION AY297438
  VERSION AY297438.1 GI:34148150
  KEYWORDS
  SOURCE Aiptasia pulchella
  ORGANISM Aiptasia pulchella
  Eukaryota; Metazoa; Cnidaria; Anthozoa; Zoantharia; Actiniaria;
  Aiptasiidae; Aiptasia.
  REFERENCE
  AUTHORS Lin.C., Tsai.J., Chen.M., Wang.L. and Fang.L.
  TITLE Characterization of 5'-flanking region of the anemone ATP
  riboylation factor 1 and actin genes
  JOURNAL Unpublished
  REFERENCE
  AUTHORS Lin.C., Tsai.J., Chen.M., Wang.L. and Fang.L.
  TITLE Direct Submission
  JOURNAL Submitted (13-MAY-2003) Research & Planning, National Museum of
  Marine Biology & Aquarium, 2 Houwan Road, Checheng, Pintung 944,
  Taiwan
  FEATURES             source
    Location/Qualifiers
      1..1440
      /organism="Aiptasia pulchella"
      /mol_type="mRNA"
      /db_xref="taxon:12924"
      /note="authority: Aiptasia pulchella Carlgren 1943"
      74..1198
      /codon_start=1
      /product="beta actin"
      /protein_id="AAQ62633.1"
      /db_xref="GI:34148151"
      /translation="MMMLPLLFLDNGSGMCKGFGAGDAPRAVPSIVGRPRHQGMV
      GMGKDSYVGDEAQSKRGILTLKYEIHGIVTNDMEKIWHHTFVNSLRVAPEDHPV
      LITEAPLPAKREKWTQIMETENSAMVTAIQAVLSYASGRTTGIIVFDGCVSH
      TVPIYEGALPHAIRLDLAGDLTDYLMKILTERGYSTTAAEIRVDLKEKLCYV
      AUDFEQMOTAASSSSSELYELPGQVITLNERFCPEALPQSFGLMGESAGHET
      TYSIMKCDADIRKDIYANTVLSGGSTMPFGIADRMQKEITSLAPTWKIKLIAPPER
      KYSVMIGGSILASLSTFQOMWISQKOEYDESGPSIVHRKCF"
  ORIGIN
    Query Match      1.3%; Score 49; DB 3; Length 1440;
    Best Local Similarity 100.0%; Pred. No. 5.8e-13;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 3714 AAATGAAAAA... 1440 bp mRNA linear PRI 21-OCT-2003
  Db 1388 AAATGAAAAA... 1436
  RESULT 49
  LOCUS BC042195 1499 bp mRNA linear PRI 21-OCT-2003
  DEFINITION Homo sapiens malonyl-CoA:acyl carrier protein transacylase
  (malonyltransferase), mRNA (cdna clone MGC:47838 IMAGE:6052380),
  complete cds.
  ACCESSION BC042195
  VERSION BC042195.1 GI:27502988
  KEYWORDS MGC.
  SOURCE Homo sapiens (human)
  ORGANISM Homo sapiens
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  REFERENCE
  AUTHORS Strausberg.R.L., Feingold.E.A., Grouse.L.H., Derge.J.G.,
  Klausner.R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
  Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat.N.K.,
  Hopkins.R.F., Jordan.H., Moore.T., Max.S.I., Wang,J., Hsieh.F.,
  Diatchenko.L., Marusina.K., Farmer.A.A., Rubin.G.M., Hong.L.,

```


Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A.C., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalys, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

12477932

2 (bases 1 to 1499)

Strausberg, R.

Direct Submission

Submitted (02-JAN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nrci.nih.gov

Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Mastello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgueon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>

Series: IRAX Plate: 82 Row: 0 Column: 7

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27477044.

Location/Qualifiers

1. .1499

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:47838 IMAGE:5052380"

/tissue_type="Testis, embryonal carcinoma"

/clone_lib="NIH MGC 92"

/lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1. .1499

/gene="MT"

/note="Synonyms: BK1191B2.3, MGC47838"

/db_xref="LocusID:27349"

214. .1215

/codon_start=1

/product="malonyl-CoA:acyl carrier protein transacylase (malonyltransferase)"

/protein_id="AAH42195.1"

/db_xref="GI:27502989"

/db_xref="LocusID:27349"

/translation="MPGQSVLLFPQGSQVVMGRGLNYPVRELYAARRVLGYD LLELSLHGQETLDRVHCQPAIFVLAALAEKHLHLPSPFIENCAVAAAGFVGFEFA

LVPAGAMEFAGLYAVKIRAEAMQSEAVPSGMLSVLQPOSKFNACLEARHCKSLGIENPVCVSNLPPDCRVISGHQEARLRFQKSSKFHRETRMLPVSGAFHTRIMEPAVELTQALKAVIDIKKLVSVYSNVGHRYEHPGCHIKKLAQQLVSPVKWEQTHAIYERKKGSGFPQTFEVPGRQLGAILKSCNQAWKSYSADVLUQTLHVDLDPQEPFR" 232. .1128

misc_feature

3714 AAATGAAAAA 1.3%; Score 49; DB 9; Length 1499; Best Local Similarity 100.0%; Pred. No. 5.8e-13; Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0; Qy 3714 AAATGAAAAA 3762 Db 1370 AAATGAAAAA 1418

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1499; Best Local Similarity 100.0%; Pred. No. 5.8e-13; Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA 3762 Db 1370 AAATGAAAAA 1418

RESULT 50

BC008836

LOCUS

DEFINITION

Homo sapiens transient receptor potential cation channel, subfamily C, member 4 associated protein, mRNA (cDNA clone IMAGE:3530893), partial cds.

ACCESSION

BC008836

VERSION

BC008836.2 GI:39645210

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1595)

AUTHORS

Klausner, R.D., Collins, F.S., Wagner, L.H., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Sapletenko, L., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Schetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahy, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalys, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1595)

Strausberg, R.

Direct Submission

Submitted (25-MAY-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

On Dec 9, 2003 this sequence version replaced gi:14250733.

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: Rubin Laboratory

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mc@hgrl.nih.gov
 Akhter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
 Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
 Dietrich N.L., Granite S., Guan X., Gupta K., Haghighi P.,
 Hansen N., Ho S.-L., Karlins E., Kwong P., Laric P., Legaspi R.,
 Maduro Q.L., Masello C., Maskeri B., Mastrian S.D., McCloskey J.C.,
 McDowell J., Pearson R., Stantropop S., Thomas P.J., Touchman J.W.,
 Turgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
 Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 20 Row: a Column: 9

FEATURES
 source
 1. 1595
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3530893"
 /tissue_type="Muscle, rhabdomyosarcoma"
 /clone_lib="NIH MGC 17"
 /lab_host="DH10B-R"
 /note="vector: pOTB7"
 <1. 1595
 /gene="TRPC4AP"
 /notes="synonyms: DKFZP727M231, DKFZp586C1223, dJ756N5.2,
 TRP4AP, C20orf188, TRUSS"
 /db_xref="LocusID:26133"
 <1. 794
 /gene="TRPC4AP"
 /codon_start=3
 /product="TRPC4AP protein"
 /protein_id="AAH08836.2"
 /db_xref="GI:39645211"
 /db_xref="LocusID:26133"
 /translations="QARAVSFLRGTTSYADQFLKRLGELHLYCIVDSECKSRDV
 LQSFVDLGLMFLVSECELLAYISQVPTQMSFLRLNIHVLTLQENVCINTSLD
 RFENQDMKAEVLSCELLAYISQVPTQMSFLRLNIHVLTLQENVCINTSLD
 ILMARKERLPLVRLRLQMERKSKYPGLNPNHNLAPFQWQHYLHKDKOSTCLEN
 SSCISFSYKWTVSILLNPRQSPSALVSIEPPYMDIDRDFTEE"

gene
 3714 AAATGAAAAA... 1655 bp mRNA linear PRI 19-NOV-2003
 1531 AAATGAAAAA... 1655 bp mRNA linear PRI 19-NOV-2003

CDS
 3714 AAATGAAAAA... 1655 bp mRNA linear PRI 19-NOV-2003
 1531 AAATGAAAAA... 1655 bp mRNA linear PRI 19-NOV-2003

ORIGIN
 Query Match 1.3%; Score 49; DB 9; Length 1595;
 Best Local Similarity 100.0%; Pred. No. 5.7e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 51
 BC039725
 LOCUS Homo sapiens cdna clone MGC:47828 IMAGE:5227869, complete cds.
 DEFINITION
 ACCESSION BC039725
 VERSION BC039725.1 GI:24980796
 KEYWORDS MGC
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 1 (bases 1 to 1655)
 Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,

Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickinson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butcherfield, J.S., Krzywinski, M.I., Skalska, U., Smallos, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 1655)
 Strausberg, R.
 Direct Submission
 Submitted (12-NOV-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: http://mgc.nci.nih.gov
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue procurement: Life Technologies, Inc.
 CDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Sequencing Center (NISC)
 Gaithersburg, Maryland;
 Web site: http://www.nisc.nih.gov/
 Contact: nisc.mc@hgrl.nih.gov
 Akhter N., Ayele K., Beckstrom-Sternberg S.M., Benjamin B.,
 Blakesley R.W., Bouffard G.G., Breen K., Brinkley C., Brooks S.,
 Dietrich N.L., Granite S., Guan X., Gupta K., Haghighi P.,
 Hansen N., Ho S.-L., Karlins E., Kwong P., Laric P., Legaspi R.,
 Maduro Q.L., Masello C., Maskeri B., Mastrian S.D., McCloskey J.C.,
 McDowell J., Pearson R., Stantropop S., Thomas P.J., Touchman J.W.,
 Turgeon C., Vogt J.L., Walker M.A., Wetherby K.D., Wiggins L.,
 Young A., Zhang L.-H. and Green E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
 Series: IRAL Plate: 82 Row: f Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, GenesScan gene prediction.

FEATURES
 source
 1. 1655
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:47828 IMAGE:5227869"
 /tissue_type="Pancreas, Spleen, adult pooled"
 /clone_lib="NIH MGC_120"
 /lab_host="DH105"
 /note="vector: pCMV-SPORT6"
 191. 1162
 /codon_start=1
 /product="Unknown (protein for MGC:47828)"
 /protein_id="AAH39725.1"
 /db_xref="GI:24980797"
 /translation="MQWALAVLAFSPASQKSNLEGRKTSVIRQTGSSAETCDLA
 EGSNGYIHWYHQEGKAPQKLYDSYNSKVLESQVSGFGKTYITASTNNRLILRN
 LIENDFGVYCACTGDPYKLFQSGTTLVTDKQDADSPKFTIFLPSATKLLQK
 AGTYICLLEKFPFDIIKIHWQKSNLTILGQENKNTNDYTKMFKSWLTVPEESLDK
 EHRCTVLRHNKNKIDQIIIPPIKTDTVTDPKDSYKSDANDVITMDPKDNWSKDN
 DTLLQLNTSAYTYLLLLKSVVYFAITCCLLRRTAFCCNGEKS"
 284. 592
 /note="IGV; Region: Immunoglobulin domain variable region
 (v) subfamily"
 /db_xref="CDD:cd00099"
 611. 901
 /note="IGC; Region: Immunoglobulin domain constant region

RESULT 53					
AX188231					
LOCUS	AX188231	1676 bp	DNA	linear	PAT 06-AUG-2001
DEFINITION	Sequence	3926	from Patent	WO0142467.	
ACCESSION	AX188231				

REMARK	USA
COMMENT	<p>NIH-MGC Project URL: http://mgc.nci.nih.gov Contact: MGC help desk Email: cgapbs-remail.nih.gov Tissue Procurement: Life Technologies, Inc. cDNA Library Preparation: Life Technologies, Inc. DNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland; Contact: www.nisc.nih.gov/ Web site: http://www.nisc.nih.gov/ Contact: nisc_mgc@nhgri.nih.gov Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, F., Latic, P., Legaspi, R., Maduro, Q.L., Masetello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantiripop, S., Thomas, P.J., Touchman, J.W.,</p>

VERSION AX189231.1 GI:15139704
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
 AUTHORS Schlegel, R., Deeds, J., Berger, A. and Zhao, X.
 TITLE Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer
 JOURNAL Patent: WO 0142467-A 3926 14-JUN-2001; Millennium Predictive Medicine, Inc. (US)

FEATURES source
 1..1676
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN
 Query Match 1.3%; Score 49; DB 6; Length 1676;
 Best Local Similarity 100.0%; Pred. No. 5.7e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 1687 bp mRNA linear ROD 06-OCT-2003
 IMAGE:4911163, complete cds.

Db 1295 AAATGAAAAA... 1687 bp mRNA linear ROD 06-OCT-2003
 IMAGE:4911163, complete cds.

RESULT 54
 BC035953
 LOCUS Mus musculus RIKEN cDNA 3010015K02 gene, mRNA (CDNA clone MGC:31692
 DEFINITION IMAGE:4911163), complete cds.

ACCESSION BC035953
 VERSION BC035953.1 GI:23271138
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T.I., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stachelson, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Uesdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalek, U., Smalls, D.E., Scherch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 MEDLINE 22388257
 PUBMED 12477932

REFERENCE 2 (bases 1 to 1687)
 DIRECT SUBMISSION Strausberg, R.
 TITLE Submitted (31-JUL-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 JOURNAL NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK Contact: MGC help desk

Email: cgapbe@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 44 Row: k Column: 18
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

Location/Qualifiers
 1..1687
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="MGC:31692 IMAGE:4911163"
 /tissue type="Salivary gland, 10 week old female mouse"
 /clone_lib="NCI CGAP_SG2"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 1..1687
 /gene="3010015K02Rik"
 /db_xref="LocusID:76832"
 /db_xref="MGI:1924082"
 361..1266
 /product="3010015K02Rik protein"
 /protein_id="AAH35953.1"
 /db_xref="GI:23271139"
 /db_xref="LocusID:76832"
 /translation="MKQLIGPGQKWNNDPEERMLAATATRICAGOGGDSRREA QAGVDYFSKASVTFCKPALPMHLHFHTASRVTSSVTSQKCKPVMKKVLRK KPDGEVLVDESIVSCESTGSDGLMDLHRFMNLQFQGTESFVVTSGKFNLPCE YGIGSOEDOLICVLOREMDPPVYQDLIVASRDKFILRLDQLSNRNKGIDRVARY FEYKRDWDSMRFPGEDEHREKELRWSVRGQMLSRTEPPSKPQHVYVNNVLPTEKKRSA LRNGVRCDLNGVMPKLPFLPLSPS"

ORIGIN
 Query Match 1.3%; Score 49; DB 10; Length 1687;
 Best Local Similarity 100.0%; Pred. No. 5.7e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 1706 bp mRNA linear INV 30-NOV-2002
 Db 1614 AAATGAAAAA... 1706 bp mRNA linear INV 30-NOV-2002

RESULT 55
 AK112205
 LOCUS Clona intestinalis cDNA, clone:ciad007ml2, full insert sequence.
 DEFINITION AK112205
 ACCESSION AK112205
 VERSION AK112205.1 GI:23575098
 KEYWORDS FLI CDNA.
 SOURCE Clona intestinalis
 ORGANISM Clona intestinalis
 Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona; Phlebobranchia; Clonidae; Clona.

REFERENCE 1
 AUTHORS Satou, Y., Yamada, L., Mochizuki, Y., Takatori, N., Kawashima, T., Sasaki, A., Hamaguchi, M., Awazu, S., Yagi, K., Sasakura, Y., Nakayama, A., Ishikawa, H., Inaba, K. and Satoh, N.
 TITLE A cDNA resource from the basal chordate Clona intestinalis
 JOURNAL Genesis 33 (4), 153-154 (2002)
 MEDLINE 22191024

12203911
 PUBMED
 REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 Submitted (04-OCT-2002) Nori Satch, Kyoto University, Department of Zoology; Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
 (E-mail: satcho@ascidian.zool.kyoto-u.ac.jp, Tel: 81-75-753-4095, Fax: 81-75-705-1113)
 Ciona intestinalis cDNA Project (URL: <http://ghost.zool.kyoto-u.ac.jp/indexr1.html>).
 FEATURES
 source
 1..1706
 /organism="Ciona intestinalis"
 /mol_type="mRNA"
 /db_xref="taxon:7719"
 /clone="ciad007m12"
 ORIGIN
 Query Match 1..3%; Score 49; DB 3; Length 1706;
 Best Local Similarity 100.0%; Pred. No. 5.7e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA...AAA 3762
 DB 1583 AAATGAAAAA...AAA 1631
 RESULT 56
 BC012854
 LOCUS
 DEFINITION
 Homo sapiens actin, beta, mRNA (cDNA clone MGC:9832 IMAGE:3863361), complete cds.
 ACCESSION
 BC012854
 VERSION
 BC012854.1 GI:15277502
 KEYWORDS
 MGC.
 SOURCE
 Homo sapiens (human)
 ORGANISM
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 1729)
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G., Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D., Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L., Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., Raha S., Lequellano N.A., Peters G.J., Abramson R.D., Mullaney S.J., Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H., Richard S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., Fahey J., Hellon E., Kettman M., Madan A., Rodriguez S., Sanchez A., Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butlerfield Y.S., Krzywinski M.I., Skalska J., Smallos D.E., Scherch A., Schein J.E., Jones S.J. and Marra M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 MEDLINE
 PUBMED
 12477932
 REFERENCE
 2 (bases 1 to 1729)
 Strausberg R.
 Direct Submission
 Submitted (20-AUG-2001) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 REMARK
 CONTACT: MGC help desk
 Email: cgapsb@mail.nih.gov

Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: angbcm@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegod, H., Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 21 Row: h Column: 1
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.
 FEATURES
 source
 1..1729
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:9832 IMAGE:3863361"
 /tissue type="Ovary, adenocarcinoma"
 /clone_lib="NIH MGC_66"
 /lab_host="DH10B"
 /notes="Vector: pCMV-SPORT6"
 Gene
 1..1729
 /genes="ACTB"
 /db_xref="LocusID:60"
 /db_xref="MIM:102630"
 CDS
 3..1085
 /codon_start=1
 /product="ACTB protein"
 /protein_id="AAH12854.1"
 /db_xref="GI:15277503"
 /db_xref="LocusID:60"
 /translation="MCKAGFAGWDADPRVFPISIVGRPHQVYMGQKDSYVGDPAQ SKRGILTKYPIEHGIVTMDMEKIWHHTFYNELRVAPEEHPVLITLAPNPKANLE KWTQIMFETNPAMTVAIQALVSLYASGRTTGVIMSDGVTHTVPIYEGVALPHAI LRDLAQRDLTDYMKILTERGVSTTTAERIVRIDIKELCVALDPEQMATRASS SSLEKSYELPDQGVITIGNERPCPEALPQPSFLMGESGJIHETTFNSIMKCDVDIRK DLTFAVLSGGTITWFGIADRMKCKEITALAPSTMKIKLIAPPKYSYVWIGSILASL STFQQMWISKQYDESGPSIVHRKCP"
 misc_feature
 3..1082
 /note="actin; Region: Actin"
 /db_xref="CDD:pfam00022"
 ORIGIN
 Query Match 1..3%; Score 49; DB 9; Length 1729;
 Best Local Similarity 100.0%; Pred. No. 5.7e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA...AAA 3762
 DB 1680 AAATGAAAAA...AAA 1728
 RESULT 57
 BC061980
 LOCUS
 DEFINITION
 Rattus norvegicus cDNA clone MGC:72261 IMAGE:5597975, complete cds.
 ACCESSION
 BC061980
 VERSION
 BC061980.1 GI:38541365
 KEYWORDS
 MGC.
 SOURCE
 Rattus norvegicus (Norway rat)
 ORGANISM
 Rattus norvegicus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
 REFERENCE
 1 (bases 1 to 1737)
 Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D., Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Datchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schetz,T.E., Brownstein,M.J., Ustin,T.B., Toshiyuki,S., Carinici,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullaly,S.J., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Faley,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalhus,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 1737)

Strausberg,R.

Direct Submission

Submitted (10-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabs@mail.nih.gov

Tissue Procurement: Jeff Green/Paturu Kondaiah, NCI.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mcg@hri.nih.gov

Akter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Hagnighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Tsurgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IRAK Plate: 137 Row: b Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 31377483.

FEATURES

source

1..1737

/organism="Rattus norvegicus"

/mol_type="mRNA"

/db_xref="taxon:10116"

/clone="MGC:72261 IMAGE:5597975"

/tissue_type="Prostate, pool of NCI_CGAP_Pr30, 40, 41 and NCI_CGAP_Pr29, 39, 42"

/clone_lib="NCI CGAP_Pr49"

/lab_host="DH10B"

/note="Vector: PCMV-SF076.1"

46..828

/codon_start=1

/product="Unknown (protein for MGC:72261)"

/protein_id="AAH61980.1"

/db_xref="GI:38541366"

/translation="MAKESGASHNGPHEHLYPIAKDQNSPIELHTKDIRDPSL QWVSYPGSAKTILNGKTVVDFDFRSMRLRGGLSPYRLRQFLHWSGSD HGSEHTVGVKAAEHLVHMNPKYNTFGALKQPDIAVYFLIKRGKGEFQILL DALDKIKTKGKAPFNHFDPSCLFFACRDYWTYHGSFTTTPFCBICVILLKBPMTVS

CDS

misc_feature

SDQMAKRLSLPASAENPEVPLVGNWRPQPPIKGRVVRASPK"

58..819

/note="carb_anhydriase; Region: Eukaryotic-type carbonic anhydriase"

/db_xref="CDD:pfam00194"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 1737;

Best Local Similarity 100.0%; Pred. No. 5.7e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1786 bp mRNA linear PRI 16-AUG-2001

Db 1683 AAATGAAAAA testis cDNA clone:QtsA-14034, full insert

RESULT 58

AB070107

LOCUS

DEFINITION

Macaca fascicularis testis cDNA clone:QtsA-14034, full insert sequence.

AB070107

AB070107.1 GI:15208054

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Macaca fascicularis (crab-eating macaque)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae; Cercopithecinae; Macaca.

REFERENCE

1

Hashimoto,K., Osada,N., Hida,M., Kusuda,J., Tanuma,R., Hirai,M., Terao,K. and Sugano,S.

Isolation of novel full-length cDNA clones from macaque testis cDNA libraries

Unpublished

2 (bases 1 to 1786)

Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.

Direct Submission

Submitted (14-AUG-2001) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan

E-mail: khashi@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genebank/>, Tel: 81-3-5285-1111(ex.2120), Fax: 81-3-5285-1181)

Lab host: TOP10

Vector: pME18S-FL3 (Acc.No. AB009864)

R. Site1: DraIII (CACTGTGTG)

R. Site2: DraIII (CACTGTGTG)

Description: 1st strand cDNA was primed with an oligo(dT) primer using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., University of Tokyo, Institute of Medical Sciences).

Science: Custom primer used for sequencing

(5' end primer [CTTCTGCTCTAAAGCTCG];

3' end primer [CGACCTCGACCTCGACACA]).

Location/Qualifiers

1..1786

/organism="Macaca fascicularis"

/mol_type="mRNA"

/db_xref="taxon:9541"

/clone="QtsA-14034"

/sex="male"

/tissue_type="testis"

/clone_lib="macaque testis cDNA library QtsA"

/dev_stage="adult"

895..1164

/codon_start=1

/product="hypothetical protein"

/protein_id="BAB63052.1"

FEATURES

source

CDS

/db_xref="GI:15208055"
/translation="MPCALLETGDAPEIATIATQVFMQCFLEALGKENKOLRFALKTY
PPTSPSLAMVLLQDPQVPRGRWLQTLKHSSELLREAVDQTHG"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1786;
Best Local Similarity 100.0%; Pred. No. 5.7e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
Db 1643 AAATGAAAAA 1691

RESULT 59
AY118452 1806 bp mRNA linear INV 15-JUN-2002
LOCUS
DEFINITION
Drosophila melanogaster AT09813 full insert cDNA.

ACCESSION
AY118452.1 GI:21428321
VERSION
FLI_CDNA.
KEYWORDS
Drosophila melanogaster (fruit fly)
SOURCE
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.

REFERENCE
1 (bases 1 to 1806)
Stapleton, M., Brokstein, P., Hong, L., Agbayani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Dresnak, D., Farfan, D., Frise, E.,
George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G.,
Miranda, A., Mungall, C. J., Nunoo, J., Pacleb, J., Paragas, V., Park, S.,
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.
and Celniker, S.

TITLE
Direct Submission
Submitted (07-JUN-2002) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA

COMMENT
Sequence submitted by:
Lawrence Berkeley National Laboratory
Berkeley, CA 94720

This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu

FEATURES
source
1..1806
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
gene
1..1806
/gene="CG7804"
CDS
341..1297
/db_xref="FLYBASE:FBgn0036496"
/gene="CG7804"
/note="Longest ORF"
/codon_start=1
/product="AT09813p"
/protein_id="AA049821.1"
/db_xref="GI:21428322"
/db_xref="FLYBASE:FBgn0036496"
/translation="MVFVHVSEKSGDEPIELPAEDGTLTLLSTLQAQFPSSSGLYRN"

VDTKAVRVRNREGRLYSPSEETGWGEYHFCVFPKKNROSENLENSTAKTRTEA
HLRCPLDILVNTTEQDLREYFYGVKAEIKDTRSGHSGKGFYRFGSYDV
QMLVSKHSDIGRWCEVKGVPASGMGNQEPKVFVGRCTEDIEADRDREYFSKFEV
IDVFIPKPFRAFSPVFLDPPVPRVVCGEKHIKGSVHVSTAKNKVQNKQLPQTN
NTNLDNNFKQCPANNFRMHPANNFMSHSPNPHGYQMKRVNN"

ORIGIN
Query Match 1.3%; Score 49; DB 3; Length 1806;
Best Local Similarity 100.0%; Pred. No. 5.7e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
Db 1758 AAATGAAAAA 1806

RESULT 60
BC029874 1835 bp mRNA linear PRI 07-OCT-2003
LOCUS
DEFINITION
Homo sapiens protein inhibitor of activated STAT protein PIASy,
mRNA (cDNA clone MGC:35296 IMAGE:5176540), complete cds.

ACCESSION
BC029874
VERSION
BC029874.1 GI:20987516
KEYWORDS
MGC.
SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
1 (bases 1 to 1835)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Caeavant, T.L.,
Schett, T.B., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carrincci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shcherenko, Y.,
Bouford, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallus, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

JOURNAL
MEDLINE
22386257
PUBMED
12477932
REFERENCE
2 (bases 1 to 1835)
Strausberg, R.
Direct Submission
TITLE
Submitted (06-MAY-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

REMARK
COMMENT
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC
Web site: http://www.hgsc.bcm.tmc.edu/cdna/
Contact: amg@bcm.tmc.edu
Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulsegad, H.,
Kowis, C.R., Speed, A.J., Martin, R.G., Muzny, D.M., Nanavati,
A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 51 Rows: m Column: 22.
Location/Qualifiers: 22.

FEATURES

source

1. 1835
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:35296 IMAGE:5176540"
/tissue_type="Brain, Lung, Testis, adult, pooled whole"
/clone_lib="NIH MGC_115"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1835
/gene="FIASy"
/note="synonym: MGC35296"
/db_xref="LocusID:51588"
/db_xref="MIM:605989"
11. 1543
/codon_start=1
/product="protein inhibitor of activated STAT protein FIASy"
/protein_id="AAH29874.1"
/db_xref="GI:20987517"
/db_xref="LocusID:51588"
/translation="MAELVAKNMVSRVSDQLMLGFGVGRSGKGLKHELVTRALQ
LVQFSPLEPKKIKELVETRYAKKNVSPAPQHPRLDPLMTHTSYDRAGAVPTPLA
GNIDYPLVYKXINGLRLPAKTLKPEVRLVLPFFNMDELKPTLVPPNNKELQ
ESFCIFLTPROVLEIRSLRLOPGVAVQVLRICYSIDTSCPDQYFNIAVKNH
SYCSVFGYIPNKGVERPCRPRLNTHLMTLSATNRITVWNGVSKSVLYLV
RQLTSSELLRLKTIQVHPELKVKEKLRLDPDSIATTVGRVSLICPLVKNRLS
VFCRAETCAHLCQFPDVFYLVQMKKPTWPCVCDKAPYQDLIIDGLLSKISECD
ADEIVLVGSGCPTRAERKSCSPQGAIVLIPSDANGLLPAPSVNGSGALGTGGG
GPGVSMNGKPGADVLDLTLDSSSEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
AC"

gene

CDS

11. 1543
/product="protein inhibitor of activated STAT protein FIASy"
/protein_id="AAH29874.1"
/db_xref="GI:20987517"
/db_xref="LocusID:51588"
/translation="MAELVAKNMVSRVSDQLMLGFGVGRSGKGLKHELVTRALQ
LVQFSPLEPKKIKELVETRYAKKNVSPAPQHPRLDPLMTHTSYDRAGAVPTPLA
GNIDYPLVYKXINGLRLPAKTLKPEVRLVLPFFNMDELKPTLVPPNNKELQ
ESFCIFLTPROVLEIRSLRLOPGVAVQVLRICYSIDTSCPDQYFNIAVKNH
SYCSVFGYIPNKGVERPCRPRLNTHLMTLSATNRITVWNGVSKSVLYLV
RQLTSSELLRLKTIQVHPELKVKEKLRLDPDSIATTVGRVSLICPLVKNRLS
VFCRAETCAHLCQFPDVFYLVQMKKPTWPCVCDKAPYQDLIIDGLLSKISECD
ADEIVLVGSGCPTRAERKSCSPQGAIVLIPSDANGLLPAPSVNGSGALGTGGG
GPGVSMNGKPGADVLDLTLDSSSEDEDEDEDEDEDEDEDEDEDEDEDEDEDEDE
AC"

misc_feature

44. 148
/note="SAP; Region: Putative DNA-binding (bhelical) motif predicted to be involved in chromosomal organisation"
/db_xref="CDD:smart00513"
974. 1132
/note="zf-MIZ; Region: pfam02891, zf-MIZ, MIZ zinc finger"
/db_xref="CDD:pfam02891"

misc_feature

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1835;
Best Local Similarity 100.0%; Pred. NO. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAA 3762

Db 1732 AATGAAA 1780

RESULT 61

BC061320 1838 bp mRNA linear VRT 04-NOV-2003
LOCUS Silurana tropicalis cDNA clone MGC:75804 IMAGE:5382264, complete

DEFINITION

cds.
BC061320 1838 bp mRNA linear VRT 04-NOV-2003
LOCUS Silurana tropicalis cDNA clone MGC:75804 IMAGE:5382264, complete

ACCESSION

BC061320.1 GI:38174055

VERSION

MGC.

SOURCE

Silurana tropicalis (Xenopus tropicalis)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Xenopodinae; Silurana.

1 (bases 1 to 1838)

Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.

and Richardson, P.

Genetic and genomic tools for Xenopus research: The NIH Xenopus

initiative

Dev. Dyn. 225 (4), 384-391 (2002)

JOURNAL

22341132

MEDLINE

12454917

PUBMED

REFERENCE

AUTHORS

2 (bases 1 to 1838)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stachewicz, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Morley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smallos, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257

MEDLINE

12477932

PUBMED

REFERENCE

AUTHORS

3 (bases 1 to 1838)
Strausberg, R.
Direct Submission
Submitted (03-NOV-2003) National Institutes of Health, Mammalian Gene Collection (MGC) Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT

NIH-MGC project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: gcaps-x@mail.nih.gov
Tissue Procurement: Robert M. Grainger
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing: The I.M.A.G.E. Consortium (LLNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: <http://www-shgc.stanford.edu>
Contact: (Dickson, Mark) mdc@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers, R. M.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 143 Row: g Column: 7
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis. Similarity but not identity to protein.

Location/Qualifiers
1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

FEATURES

source

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

FEATURES

source

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"

1. 1838
/organism="Silurana tropicalis"
/mol_type="mRNA"
/db_xref="taxon:8364"
/clone="MGC:75804 IMAGE:5382264"
/tissue_type="Embryo, Silurana tropicalis, neurula (stages 14-19)"
/clone_lib="NICHD_XCC_Emb6"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6.1"
128. 1000
/codon_start=1
/product="Unknown (protein for MGC:75804)"
/protein_id="AAH61320.1"
/db_xref="GI:38174055"


```

misc_feature
NLKETQGTQNTAHS"
578..859
/note="THUMP; Region: THUMP domain. The THUMP domain is
named after thionuridine synthases, methylases and
PSUs. The THUMP domain consists of about 110 amino acid
residues. It is predicted that this domain is an
RNA-binding domain that adopts an alpha/beta fold similar
to that found in the C-terminal domain of translation
initiation factor 3 and ribosomal protein S8. The THUMP
domain probably functions by delivering a variety of RNA
modification enzymes to their targets"
/db_xref="CDD:pfam02926"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 1838;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1778 AATGAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1826

RESULT 62
BC044938 1850 bp mRNA linear PRI 01-DEC-2003
LOCUS Homo sapiens cDNA clone IMAGE:5190011, complete cds.
DEFINITION BC044938
ACCESSION BC044938.1 GI:27924313
VERSION
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE
1 (bases 1 to 1850)
Strausberg R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.P., Zesberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldi, M.P., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raja, S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettaman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 1850)
Strausberg R.
Direct Submission
Submitted (23-JAN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NHL-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cagbs-r@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca

```

```

Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandon, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline
Schein, Duane Smilau, Michael Smith, Lorraine Spencer, Jeff Stott,
Michael Thorne, Miranada Teai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 87 Row: j Column: 14
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 1630585
This clone has the following problem: The cds is short compared to
the longest cds in the locus.

FEATURES
Location/Qualifiers
1..1850
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5190011"
/tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
/clone_lib="NIH MGC_116"
/lab_host="DH10B"
/notes="Vector: PCMV-SPORT6"
371..583
/codon_start=1
/product="Unknown (protein for IMAGE:5190011)"
/protein_id="AAH4938.1"
/db_xref="GI:27924314"
/translation="MNVSSSHFVSALTVPFNGSKLSIKIEDTPVDDPSLKILVANN
DTLLKMSCHPVSSDDGLHFLKLE"

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 1850;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1735 AATGAGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1793

RESULT 63
AB063046 1894 bp mRNA linear PRI 13-JUN-2001
LOCUS Macaca fascicularis brain cDNA clone:QMOA-11972, full insert
sequence.
DEFINITION AB063046
ACCESSION AB063046.1 GI:14388491
VERSION
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopithecinae; Macaca.
REFERENCE
1 (sites)
Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,
Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
Isolation of full-length cDNA clones from macaque brain cDNA
libraries
Unpublished
2 (bases 1 to 1894)
Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
Direct Submission
Submitted (11-JUN-2001) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
Tel: 81-3-5285-1111 (ex.2120), Fax: 81-3-5285-1181)
Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)

```

R. Site1: DraIII (CACTGTGTG)
 R. Site2: DraIII (CAACATGTG)
 Description: 1st strand cDNA was primed with an oligo(dT) primer
 [ATGCGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
 using specific 5' and 3' primers and amplified by PCR. The PCR
 product was digested with SfiI and size selection was performed to
 exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
 into distinct DraIII sites of pME18S-FL3. XhoI sites just outside
 the DraIII sites can be used to isolate the cDNA insert. Libraries
 were constructed by oligo-capping method
 (Sugano et al., , Institute of Medical Science, University of
 Tokyo).

Custom primer used for sequencing
 (5' end primer [CTTCTGCTCTAAAGCTGCG];
 3' end primer [CGACTGCTGAGCTGCACCA]).

FEATURES

source

Location/Qualifiers
 1..1894
 /organism="Macaca fascicularis"
 /mol_type="mRNA"
 /db_xref="taxon:9541"
 /clone="QmoA-11972"
 /sex="male"
 /tissue types="medulla oblongata"
 /clone_lib="macaque brain cDNA library QmoA"
 /dev_stage="adult"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1894;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGCAA 3762
 DB 1773 AATGCAA 1821

RESULT 64

AR300635 1933 bp DNA linear PAT 12-JUN-2003
 LOCUS AR300635 Sequence 3 from patent US 6537811.
 DEFINITION AR300635
 ACCESSION AR300635
 VERSION AR300635.1 GI:31688184
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1933)
 AUTHORS Freier,S.M.
 TITLE Antisense inhibition of SAP-1 expression
 JOURNAL Patent: US 6537811-A 3 25-MAR-2003;

FEATURES
 source
 1..1933
 /organism="unknown"
 /mol_type="genomic DNA"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 1933;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGCAA 3762
 DB 1878 AATGCAA 1926

RESULT 65

AX146886 1933 bp DNA linear PAT 08-JUN-2001
 LOCUS AX146886 Sequence 20 from Patent WO0136674.
 DEFINITION AX146886
 ACCESSION AX146886
 VERSION AX146886.1 GI:14346186
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Bull,J.H., Ellison G. and Paskins,L.D.

Diagnostic methods for the detection of prostate disorders

JOURNAL

Patent: WO 0136674-A 20 25-MAY-2001;

FEATURES

source

Location/Qualifiers

1..1933
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 1933;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGCAA 3762
 DB 1878 AATGCAA 1926

RESULT 66

HUMSAPIA 1933 bp mRNA linear PRI 17-DEC-1993
 LOCUS HUMSAPIA Homo sapiens/SRP accessory protein 1A (SAP-1) mRNA, complete cds.
 DEFINITION M85165
 ACCESSION M85165.1 GI:429185
 VERSION serum response factor; SAP-1; Elk-1.
 KEYWORDS
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Dalton,S. and Treisman,R.
 Characterization of SAP-1, a protein recruited by serum response
 factor to the c-fos serum response element
 Cell 68 (3), 597-612 (1992)
 92154673
 1339307
 2 (bases 1 to 1933)
 Treisman,R.
 Direct Submission
 Submitted (05-MAR-1992) Richard Treisman, Transcription Laboratory,
 Imperial Cancer Research Fund, London, England
 On Nov 29, 1993 this sequence version replaced gi:338036.
 Original source text: Homo sapiens cDNA to mRNA.

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

FUBMED

REFERENCE

AUTHORS

JOURNAL

COMMENT

FEATURES

source

Location/Qualifiers

1..1933
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"

gene

CDS

1..1933
 /gene="SAP-1"
 /gene="SAP-1"
 /notes="Homology region A with Elk-1 protein (Ets domain)
 is bp 150-417, amino acids 1-89; Homology region B with
 Elk-1 protein, required for cooperative ternary complex
 formation with SRF is bp 565-617, amino acids 136-157;
 Sequence diverges from SAP-1B at bp 1229, amino acid 360;
 Homology region C with Elk-1 protein, core of regulated
 transcription activation domain, is bp 1203-1355, amino
 acids 352-402; (S/T)P motifs conserved between SAP-1A and
 Elk-1 are located at amino acids T354, T361, T366, S381,
 S387, T420, S425, corresponding to bp 1209-1214, bp
 1230-1235, bp 1245-1250, bp 1290-1295, bp 1308-1313, bp
 1407-1412, bp-1422-1427"

codon_start=1

/product="SAP-1A protein"

/protein_id="AAA03631.1"

```

/db_xref="GI:429186"
/translation="WDSALITVWFLLOLQKPNKHMICWTSNDGQFKLQAEVARL
WGIRKNPNMYDKLSALRYVYVNIKKVNGKQFYKVSPEILNMDPMTVGRIE
GDCSLNFSVSSKGVNGKQKPPQGAKTSSRNDYHSLGYSFTLNSLNSNV
KFLXIKTENPAELKARKSPQETPSPVIRKPPVPEVPAATISIGPSISPS
KETQALTVSKPLSLEAPTSASNVPATPPPISSIPPIQEPRTSPPLSHHP
DIDIDIVASQPMELPENTLESKQDQSVLLEKDNNSRKKPKKGLGLAFTLVIT
SSDSPGLILSPISPTASLTAPFQSQTPIILTPSLSSIHFWSTLSPLSPARLQ
GANTLFPFPSVLNHSHPGPTLSGLDGPSTGPGFSPDLQKT"

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 1933;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
Db 1878 AAATGAAAAA 1926

RESULT 67
BC056597
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 1940)
Straussberg,R.U., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.F., Schuler,G.D.,
Altshul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.M., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Heieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schneitz,T.E., Brownstein,M.J., Uedin,T.B., Tothiyuki,S.,
Carninci,P., Frange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahay,J., Helton,E., Kertanen,M., Madan,A.C., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Boutfard,G.G., Blakesley,R.W., Touchman,J.W., Green,B.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalusz,D.E.,
Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22386257
JOURNAL
MEDLINE
PUBMED
12477932
REFERENCE
2 (bases 1 to 1940)
Straussberg,R.
Direct Submission
Submitted (15-AUG-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgs.nci.nih.gov
Contact: MGC help desk
Email: gcgabs-remail.nih.gov
Tissue Procurement: Dr. Chi-Bin Chien
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/

```

Contact: nisc_mgs@nih.gov
 Akhter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
 Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
 Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
 Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
 Maduro,Q.L., Mastello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
 McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
 Tsugeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
 Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Place: 121 Row: n Column: 11
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis, Similarity but not identity to protein.

FEATURES
 source

Location/Qualifiers
 1..1940
 /organism="Danio rerio"
 /mol_type="mRNA"
 /db_xref="taxon:7955"
 /clone="MGC:65863 IMAGE:6800154"
 /tissue_type="Embryo, zebrafish whole embryo 72 hpf"
 /clone_lib="NCI CGAP_ZEmb3"
 /lab_host="DH10B"
 /note="Vector: PCMV-SPORT6.1"
 270..1196

CDS

/codon_start=1
 /product="Unknown (protein for MGC:65863)"
 /protein_id="AAH56597.1"
 /db_xref="GI:33989472"
 /translation="MSTKMTTFYDSDLSNAPSOHDSATFGYNHKKHNTLNLSDP
 AGNLKPLRAKASDILTPDVLKLAPELERLIQSSNGMITTPPTQFLCPKPNV
 TDEQGPAREGFVRAALAEHLHQHMVNTSAPOTTINSSMAPVSSIAGGAVSSAMEADP
 PYADINFPNPAISSANPAWSPFAPQPFVQHPRLQALKKEPQVPEWPGETPP
 LFDIMESQERIKAEKRNRIKASCKRKLRIKLEDKVKTLKSCNSELASTAN
 MLREQVALQKQKVMHNSGCOLMLTQQLQTF"
 282..923

misc_feature

/note="JNK; Region: Jun-like transcription factor. The
 c-Jun NH(2)-terminal kinase (JNK) is a member of an
 evolutionarily conserved sub-family of mitogen-activated
 protein (MAP) kinases"
 /db_xref="CDD:pfam03957"
 948..1139

misc_feature

/note="bZIP; Region: bZIP transcription factor. The Pfam
 entry includes the basic region and the leucine zipper
 region"
 /db_xref="CDD:pfam00170"

ORIGIN

Query Match 1.3%; Score 49; DB 5; Length 1940;
 Best Local Similarity 100.0%; Pred. No. 5.6e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

Db 1850 AAATGAAAAA 1898

RESULT 68

BC051026
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 1970)
 Straussberg,R.
 Direct Submission

REFERENCE

1 (bases 1 to 1970)
 Straussberg,R.
 Direct Submission

REMARK
 COMMENT

JOURNAL

Submitted (11-APR-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT

Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Ahter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B., Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S., Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P., Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R., Maduro,Q.L., Masiello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C., McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W., Turgeon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L., Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>
Series: IRAC Plate: 93 Row: 1 Column: 17
This clone was selected for full length sequencing because it passed the following selection criteria: Genomescan gene prediction.

FEATURES

source Location/Qualifiers

1. 1970
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5202626"
/tissue_type="lung, spleen, fetal, pooled"
/clone_lib="NIH MGC_122"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1970;

Best Local Similarity 100.0%; Pred. No. 5.6e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

Db 1880 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1928

RESULT 69

AB060842

LOCUS

AB060842 1988 bp mRNA linear PRI 13-JUN-2001

DEFINITION

Macaca fascicularis brain cDNA clone:QtrA-10930, full insert

ACCESSION

AB060842

VERSION

AB060842.1

KEYWORDS

oligo capping; fis (full insert sequence).

SOURCE

Macaca fascicularis

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;

Cercopithecinae; Macaca.

REFERENCE

1 (sites)

Osada,N., Hida,M., Kusuda,J., Tanuma,R., Iseki,K., Hirai,M.,

Terao,K., Suzuki,Y., Sugano,S. and Hashimoto,K.

Isolation of full-length cDNA clones from macaque brain cDNA

libraries

Unpublished

2 (bases 1 to 1988)

Hashimoto,K., Osada,N., Hida,M., Kusuda,J. and Sugano,S.

Direct Submission

JOURNAL

Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of Infectious Diseases, Division of Genetic Resources; 23-1, Toyama 1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genebank/>, Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)

COMMENT

Lab host: TOP10
Vector: pME18S-FL3 (Acc.No. AB009864)
R. Site1: DraIII (CACTGCTG)
R. Site2: DraIII (CACCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGGGCTTTT...TTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., Institute of Medical Science, University of Tokyo).

Custom primer used for sequencing
(5' end primer [CTTCTGCTCTAAAGCTGCG];
3' end primer [CGACCTGCACCTCGACACA]).

FEATURES

source Location/Qualifiers

1. 1988
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="QtrA-10930"
/sex="male"
/tissue_type="temporal lobe right"
/clone_lib="macaque brain cDNA library QtrA"
/dev_stage="adult"

ORIGIN

Query Match 1.3%; Score 49; DB 9; Length 1988;

Best Local Similarity 100.0%; Pred. No. 5.6e-13;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

Db 1928 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1976

RESULT 70

AF019117

LOCUS

AF019117 2015 bp mRNA linear INV 27-NOV-1997

DEFINITION

Aedes aegypti transferrin precursor, mRNA, complete cds.

ACCESSION

AF019117

VERSION

AF019117.1

KEYWORDS

GI:2645496

SOURCE

Aedes aegypti

ORGANISM

Aedes aegypti

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Nematocera; Culicoidea; Aedes;

Stegomyia.

REFERENCE

1 (bases 442 to 1971)

Beerntsen,B.T., Severson,D.W. and Christensen,B.M.

Aedes aegypti: characterization of a hemolymph polypeptide

expressed during melanotic encapsulation of filarial worms

Exp. Parasitol. 78, 313-321 (1994)

JOURNAL

2 (bases 1 to 2015)

Yoshiga,T., Hernandez,V.P., Fallon,A.M. and Law,J.H.

Mosquito transferrin, an acute-phase protein that is up-regulated

upon infection

Proc. Natl. Acad. Sci. U.S.A. 94 (23), 12337-12342 (1997)

MEDLINE

98024132

PUBMED

9356450

REFERENCE

3 (bases 1 to 2015)

Yoshiga,T.

Direct Submission

TITLE

Submitted (12-AUG-1997) Biochemistry, University of Arizona,

Tucson, AZ 85721, USA

JOURNAL

Location/Qualifiers

FEATURES

```
source
1. 2015
/organism="Aedes aegypti"
/mol_type="mRNA"
/strain="Rockefeller"
/db_xref="taxon:7159"
/dev_stages="larva"
4. 1905
/functions="iron-binding protein"
/functions="iron-transport protein"
/codon_start=1
/product="transferrin precursor"
/protein_id="AAB87414.1"
/db_xref="GI:2645497"
/translation="MTVSNALLSVVAVILLIGSQGQKTFKLPVQIMDACQDLM
AKPAAIOVQICAGRMECEKVKAREADVAVDPMVAVHMANODSVTEFT
LEPKAEPRYEGIIILVRKSDNFRSLADLRGKSKCHTGVNGVYKIPITKLKAGVFK
LATSELSPLKELKGLSLDFGSACLVGKSPNDEVNRLKRVSNLCALCEPVCVD
YDKSGTGDGALICLVNNGDVATKIYVNYKFLGVPGGAPALNPARTEDYV
ICEGSTRPITGPACSAQRPWQGTMGNDINSRFLQRLQOYQDANKNSADTKA
LKMVDRKNVLVDREVVPQGDHLNRAQYDIERDPFQNKIKLVTSLIENKCEV
MKAAYSRDVPAFECVMKSGSCVEAVEGEADVVLKGEDQDPATSKSLFEK
YDDNDVMAVADKQITRDQVLKAPLEFADSPRAVSAALFNDKQKQACPNKLASQ
GAPRIVSAGADKTRIGANKRLICPNLEQALGNVQACNVDTYMTPTGYVRKDTGQLE
DNIAHAFVALSDKPGHGAKEVVEPMEGFEFAGEKNVLFHDBRAKFGVAGNKIGEDP
ANIRKLMCL"
sig_peptide
mat_peptide
67. 1902
/product="transferrin"

ORIGIN
Query Match 1.3%; Score 49; DB 3; Length 2015;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 1966 AAATGAAAAA... 2014

RESULT 71
BC053219 2031 bp mRNA linear VRT 08-OCT-2003
LOCUS
DEFINITION
Danio rerio cDNA clone MGC:64043 IMAGE:6793489, complete cds.
ACCESSION
BC053219.1 GI:31419505
VERSION
MGC:
KEYWORDS
Danio rerio (zebrafish)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

REFERENCE
1 (bases 1 to 2031)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan B., Moore T., Max S.I., Wang J., Heide F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Scheetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Carninci P., Prange C., Raha S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullah S.J., Bosak S.A., McEwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley K.C., Hale S., Garcia A.M., Gay L.J., Rulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Hellmich M., Madan A., Young A.C., Shchegolev Y.,
Sanchez A., Whitting M., Madan A., Young A.C., Shchegolev Y.,
Bouffard G.G., Blakeley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalls D.E.,
Scherer A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

JOURNAL
MEDLINE
PUBMED
REFERENCE
2 (bases 1 to 2031)
Strausberg R.
Direct Submission
Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapps@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
cDNA Library Preparation: Invitrogen Corp
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAC Plate: 117 Row: i Column: 8
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
FEATURES
Location/Qualifiers
1. 2031
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:64043 IMAGE:6793489"
/tissue_type="Kidney, zebrafish"
/clone_lib="NCI CGAP_ZKId1"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
140..514
/codon_start=1
/product="Unknown (protein for MGC:64043)"
/protein_id="AAH53219.1"
/db_xref="GI:31419506"
/translation="MAKTDGSSFFRRSSVFMVLTSLALAFFTWTFVFRDYPVYGLG
PLGALANYVVEQHYTVYVYGFETWLIHLCEAMFALKLCSKDGIFSSARLLWFWQTF
LFGFASLGILLKYPDGRKRH"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 2031;
Best Local Similarity 100.0%; Pred. No. 5.6e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 1948 AAATGAAAAA... 1996

RESULT 72
HSU55017 2038 bp mRNA linear PRI 05-MAY-1996
LOCUS
DEFINITION
Human transketolase (TKT) mRNA, complete cds.
ACCESSION
U55017 M86521
VERSION
U55017.1 GI:1297296
KEYWORDS
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2038)
Abedinia M., Layfield R., Jones S.M., Nixon P.F. and Mattick J.S.
Nucleotide and predicted amino acid sequence of a cDNA clone
encoding part of human transketolase
```

```

JOURNAL Biochem. Biophys. Res. Commun. 183 (3), 1159-1166 (1992)
MEDLINE 92231878
PUBMED 1567394
REFERENCE 2 (bases 1 to 2038)
AUTHORS Schenk, G., Layfield, R., Candy, J.M., Duggleby, R.G. and Nixon, P.F.
TITLE Molecular evolutionary analysis of the thiamin-dependent enzyme, transketolase
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 2038)
AUTHORS Schenk, G., Candy, J.M., Duggleby, R.G. and Nixon, P.F.
TITLE Direct Submission
JOURNAL Submitted (26-MAR-1996) G. Schenk, Biochemistry, University of Queensland, Brisbane, Qld 4072, Australia
COMMENT On May 6, 1996 this sequence version replaced gi:339941.
FEATURES
    source
        1. .2038
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /chromosome="3"
            /map="3p14.3"
            /tissue_type="brain"
        1. .2038
            /gene="TKT"
        1. .1872
            /gene="TKT"
            /codon_start=1
            /product="transketolase"
            /protein_id="AAA98961.1"
            /db_xref="GI:1297297"
            /translation="MESYHKPDQKQLQALKDTANRLRISSTQATTAAGSGHPTSCCSA
            AEMNAVFFHMYKSPQDRPHNDRFVLSKGHAAPILYAVMAEAGFLAEALLNRK
            ISSLDGHPVPKQAFDTVATSGLGGLGACGMAYTKYFDKASRVYVCLLDGGLSE
            GSVWEAMAFASYKLDNLVAIDLINRLGQSDPAPLOHMDIYQRCFAFGWHAIIVDG
            HSEELKAFQAKHQPTAIATKTKGRTGVEDEKESMGKPLPKMAEQIIQBIYS
            IQSKKILATPPQEDAFSDIANIRMPISLPYKVGDKIATRAYGQALAKLGHASDR
            IIALDGTNTSTFSEIFKKEHPDRFIECYIAEQNMVSIAGVCATNRVTPFCSTFAAF
            FTRAPDQIRMAAISSEININLCSHGVSIGEDGSAQMALEDLAFERSVPTSTVFPSD
            GVTEHAAVELAANKTKICFIRTPENAIYNNEDFQVQAKVLSKDDQVTVIGA
            GVTLEHAAELALKKKNIRVLPFTIKPLDRKLIIDRSARATGRILTVEDHYIEG
            GIGEAUSSAVVGEFGITVTHLVANVRPSRGPPELLKMFIDRDAIAQAVRGLITKA"
            /translation="MESYHKPDQKQLQALKDTANRLRISSTQATTAAGSGHPTSCCSA
            AVIMADGHPVPMRYKSPQDRPHNDRFVLSKGHAAPILYAVMAEAGFLAEALLNRK
            ISSLDGHPVPKQAFDTVATSGLGGLGACGMAYTKYFDKASRVYVCLLDGGLSE
            GSVWEAMAFASYKLDNLVAIDLINRLGQSDPAPLOHMDIYQRCFAFGWHAIIVDG
            HSEELKAFQAKHQPTAIATKTKGRTGVEDEKESMGKPLPKMAEQIIQBIYS
            IQSKKILATPPQEDAFSDIANIRMPISLPYKVGDKIATRAYGQALAKLGHASDR
            IIALDGTNTSTFSEIFKKEHPDRFIECYIAEQNMVSIAGVCATNRVTPFCSTFAAF
            FTRAPDQIRMAAISSEININLCSHGVSIGEDGSAQMALEDLAFERSVPTSTVFPSD
            GVTEHAAVELAANKTKICFIRTPENAIYNNEDFQVQAKVLSKDDQVTVIGA
            GVTLEHAAELALKKKNIRVLPFTIKPLDRKLIIDRSARATGRILTVEDHYIEG
            GIGEAUSSAVVGEFGITVTHLVANVRPSRGPPELLKMFIDRDAIAQAVRGLITKA"

ORIGIN
    Query Match 1.3%; Score 49; DB 9; Length 2038;
    Best Local Similarity 100.0%; Pred. No. 5.6e-13;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 2106 bp mRNA linear PRI 09-MAY-1996
    |||||
Db 1974 AAATGAAAAA... 2022

RESULT 73
HSTRANSK HSTRANSK 2106 bp mRNA linear PRI 09-MAY-1996
LOCUS H.sapiens mRNA for transketolase.
DEFINITION X67688 S52775
ACCESSION X67688.1 GI:37266
VERSION X67688.1 GI:37266
KEYWORDS transketolase
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 2104)
AUTHORS Abedin, M., Layfield, R., Jones, S.M., Nixon, P.F. and Mattick, J.S.
TITLE Nucleotide and predicted amino acid sequence of a cDNA clone encoding part of human transketolase
JOURNAL Biochem. Biophys. Res. Commun. 183 (3), 1159-1166 (1992)
MEDLINE 92231878
PUBMED 1567394
REFERENCE 2
AUTHORS McCool, B.A., Plonk, S.G., Martin, P.R. and Singleton, C.K.

```

```

TITLE Cloning of human transketolase cDNAs and comparison of the
nucleotide sequence of the coding region in Wernicke-Korsakoff and
non-Wernicke-Korsakoff individuals
JOURNAL J. Biol. Chem. 268 (2), 1397-1404 (1993)
MEDLINE 93123263
PUBMED 8419340
REFERENCE 3 (bases 1 to 2104)
AUTHORS Singleton, C.K.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1992) C.K. Singleton, Vanderbilt University, Box
1820, Station B, Nashville, TN 37235, USA
FEATURES
    source
        1. .2106
            /organism="Homo sapiens"
            /mol_type="mRNA"
            /db_xref="taxon:9606"
            /clone="pTKa-1, pTKd-1"
            /cell_line="Hela (nucleotides residues 1-200 only)"
            /tissue_type="frontal cortex, liver"
            /clone_lib="frontal cortex and liver cDNA"
        71. .1942
            /EC_number="2.2.1.1"
            /note="homodimer"
            /codon_start=1
            /evidence="experimental"
            /product="transketolase"
            /protein_id="CAA47919.1"
            /db_xref="GI:37267"
            /db_xref="GOA:P29401"
            /translation="MESYHKPDQKQLQALKDTANRLRISSTQATTAAGSGHPTSCCSA
            AVIMADGHPVPMRYKSPQDRPHNDRFVLSKGHAAPILYAVMAEAGFLAEALLNRK
            ISSLDGHPVPKQAFDTVATSGLGGLGACGMAYTKYFDKASRVYVCLLDGGLSE
            GSVWEAMAFASYKLDNLVAIDLINRLGQSDPAPLOHMDIYQRCFAFGWHAIIVDG
            HSEELKAFQAKHQPTAIATKTKGRTGVEDEKESMGKPLPKMAEQIIQBIYS
            IQSKKILATPPQEDAFSDIANIRMPISLPYKVGDKIATRAYGQALAKLGHASDR
            IIALDGTNTSTFSEIFKKEHPDRFIECYIAEQNMVSIAGVCATNRVTPFCSTFAAF
            FTRAPDQIRMAAISSEININLCSHGVSIGEDGSAQMALEDLAFERSVPTSTVFPSD
            GVTEHAAVELAANKTKICFIRTPENAIYNNEDFQVQAKVLSKDDQVTVIGA
            GVTLEHAAELALKKKNIRVLPFTIKPLDRKLIIDRSARATGRILTVEDHYIEG
            GIGEAUSSAVVGEFGITVTHLVANVRPSRGPPELLKMFIDRDAIAQAVRGLITKA"

ORIGIN
    Query Match 1.3%; Score 49; DB 9; Length 2106;
    Best Local Similarity 100.0%; Pred. No. 5.6e-13;
    Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 2177 bp mRNA linear VRT 03-DEC-2003
    |||||
Db 2044 AAATGAAAAA... 2092

RESULT 74
BC063356 2177 bp mRNA linear VRT 03-DEC-2003
LOCUS Silurana tropicalis cDNA clone MGC:75895 IMAGE:5383371, complete
DEFINITION BC063356
ACCESSION BC063356.1 GI:38648981
VERSION BC063356.1
KEYWORDS MGC.
SOURCE Silurana tropicalis (Xenopus tropicalis)
ORGANISM Silurana tropicalis
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Silurana.
REFERENCE 1 (bases 1 to 2177)
AUTHORS Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Clifton, S.W.
and Richardson, P.
TITLE Genetic and genomic tools for Xenopus research: The NIH Xenopus
initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 2177)

```


A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 60 Row: 0 Column: 9
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

source	Location/Qualifiers
JRES	1. .2182

```

RES      Location/Qualifiers
source   1..2182
          /organism="Mus musculus"
          /mol_type="mRNA"
          /strain="FVB/N"
          /db_xref="taxon:10090"
          /clone="IMAGE:4953665"
          /issue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
          ductal carcinoma. 5 month old virgin mouse."
          /clone_lib="NCI CGAP_Mam6"
          /lab_host="DH10B"
          /note="Vector: pCMV-SPORT6"
gene      1..2182
          /gene="LOC217820"
          /db_xref="LocusID:217820"
          <1..935
          /gene="LOC217820"
          /codon_start=3
          /product="LOC217820 protein"
          /protein_id="AAH27154.1"
          /db_xref="GI:20071799"
          /db_xref="LocusID:217820"
          /translation="VFVGGKILVTRNSELIIEVGKNAACNLIVNGHVDGPIWGLATH
          PSDDPLSAEDGTVRLMDIADKXWLNKLVNHAATVCYSPGDMVALGKMGKEFI
          LLVSKLTKGKKDRRCALHDIRSPSRVLANGSSENVDPDLITGFTLNRISYCK
          DIPFVIQMFADSRGHLQVSGSGCKIKRHYEVPSGKHLVDHAALDRITWATWTILG
          EVNGVIRHAEKADVTCAKVSCHGISLVIGDFGDMGIVDFPCPEKFAHKRFLGHS
          HTVNIITGSDRHVVSGGDDCSLFWVKVCMHMPH"
          21..2176
          /gene="LOC217820"
          /note="COG2319; Region: FOG; WD40 repeat [General function:
          prediction only]"
          /db_xref="CDD:COG2319"
misc_feature

```

misc feature

ORIGIN

```
Query Match      1.3%; Score 49; DB 10; Length 2182;
Best Local Similarity 100.0%; Pred. No. 5.5e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

Qy	3714	AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	3762
Db	2130	AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	2178

RESULT 76

BC049193	BC049193	2266 bp	mRNA	linear	PRI 04-NOV-2003
LOCUS					
DEFINITION	Homo sapiens KIAA1277 protein, mRNA (cDNA clone IMAGE:445131), partial cds.				

ACCESSION BC049193

BC049193.1 GI:29351610

KEYWORDS

SOURCE	Homo sapiens (human)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

ORGANISM

REFERENCE
1 (bases 1 to 2266)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

CONTRIBUTING AUTHORS

Strausberg, R.L., Peingold, E.A., Grouse, L.H., Derge, J.G.,
Klauser, R.D., Collins, P.S., Wagner, L., Sherman, C.M., Schuler, G.D.,
Altchul, S.P., Zeebarg, B., Moore, K.H., Schaefer, C.P., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marasina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stachenko, M., Soares, M.B., Bonaldi, M.F., Casavant, T., L.
Schaeetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carinci, P., Prance, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,

Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettelman, M., Madan, A., Rodrigues, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grinwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalau, D.E.,
Schnercher, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

22388257
12477932
2 (bases 1 to 2266)
Strausberg, R.
Direct Submission
Submitted (21-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabbs-r@mail.nih.gov
Tissue Procurement: ARCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: <http://www.nisc.nih.gov/>
Contact: nisc_mgc@nhgri.nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
Dierker, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
Hansen, N., Ho, S.-I., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
Tsougeon, C., Vogt, J.-L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
Young, A.C., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IPAK plate: 93 Row: c Column: 21.

FEATURES source

Location/Qualifiers
1. .2266

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:445131"
/tissue_type="Liver, adenocarcinoma"
/clone_lib="NIH MGC 90"
```

gene

<1. .2266

```
/db xref="LocusID:27147"
```

<1. .1918
/gene="KIAA1277"

```

/codon_table2
/product="KIAA1277 protein"

```

```
/procell_id="AAH49193.1"
/db_xref="GT:29351611"
```

```
/db_xref="locusID:2/L4/"
/translation="DILPPPMKENPYE"
```

QSLSKPAFFRQNSERNFALDTRKLS
GBKTRKTPKT.VT.BTNAIYEVBRGKKRV

HSQRLVNVKSLKQAFKYPSSLARELIE
FPI.KI.FBSSF.KI.MBEAFEDOL.KATPOCF

YCRRLPGGRKRLPEVICIVSRIGCF
PEPAIGKTTIVKNFIPGSGTEVTETCF

LERRVFIADKLSILSKCCHAMVALLY


```

SSSLPLRLPLEEVLVVDLNSRFLRQMDDEDSILPRKLQVALEHILEQORNELACRQ
DEGPLDGRHGPSSPLNWEAFVFPVEIVGHVSLFTSGEREERTLOREAPRKAV
SKSLRHLEFLVEMETOMGRFIQERELRQDAKGLFEVRAQEYLLETLPSEHSGVNF
LKGLGNKKMKFLHKX"
557..847
/genre="KIAA1277"
/Note="uDENN; Region: uDENN domain. This region is always
found associated with pfam02141. It is predicted to form
an all beta domain"
/db_xref="CDD:pfam03456"
866..1330
/genre="KIAA1277"
/Note="DENN; Region: DENN (AEX-3) domain. DENN (after
differentially expressed in neoplastic vs normal cells) is
a domain which occurs in several proteins involved in Rab-
mediated processes or regulation of MAPK signalling
pathways"
1571..1774
/db_xref="CDD:pfam02141"
/genre="KIAA1277"
/Note="dbDENN; Region: dbDENN domain. This region is always
found associated with pfam02141. It is predicted to form a
globular domain. This domain is predicted to be completely
alpha helical. Although not statistically supported it has
been suggested that this domain may be similar to members
of the Rho/Rac/Cdc42 GEF family"
/db_xref="CDD:pfam03455"

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 2266;
Best Local Similarity 100.0%; Pred. No. 5.5e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 2270 bp mRNA linear PRI 18-FEB-2000
||||| 3762
Db 2201 AAATGAAAAA 2249

RESULT 77
HSM801945
LOCUS HSM801945 2270 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA; cDNA DKFZp434C2016 (from clone DKFZp434C2016).
ACCESSION AL137281
VERSION AL137281.1 GI:6807734
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 2270)
Koehler K., Beyer A., Mewes H.W., Gassenhuber J. and Wiemann S.
Direct Submission
Submitted (15-JAN-2000) MIPS, Am Klopferspitz 18a, D-82152
Martinsried, GERMANY
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email: s.wiemann@dkfz-heidelberg.de;
sequenced by BMFZ (Biomedical Research Center at the Charite,
Berlin/Germany) within the cDNA sequencing consortium of the German
Genome Project.
This clone (DKFZp434C2016) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
information about the clone and the sequencing project is available
at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
source
1..2270
location/Qualifiers
organism="Homo sapiens"
mol_type="mRNA"
db_xref="taxon:9606"
map="433.4 cR from top of Chr11 linkage group"
clone="DKFZp434C2016"
tissue_type="testis"
clone_lib="434 (synonym: htes3). Vector pSport1; host

```

```

DH10B; sites NotI + SalI"
/dev_stage="adult"
polyA_signal 2189..2194
polyA_site 2207
ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 2270;
Best Local Similarity 100.0%; Pred. No. 5.5e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 2389 bp mRNA linear ROD 12-DEC-2003
||||| 3762
Db 2209 AAATGAAAAA 2257

RESULT 78
BC064017
LOCUS BC064017 2389 bp mRNA linear ROD 12-DEC-2003
DEFINITION Mus musculus cDNA clone IMAGE:3995710, partial cds.
ACCESSION BC064017
VERSION BC064017.1 GI:39794122
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 2389)
Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L.,
Schetz T.E., Brownstein M.J., Usdin T.B., Toshiyuki S.,
Carroll P., Prange C., Raja S.S., Loquellano N.A., Peters G.J.,
Abramson R.D., Mullany S.J., Bosak S.A., McSwan P.J.,
McKernan K.J., Malek J.A., Gunaratne P.H., Richards S.,
Worley D.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahy J., Helton E., Kettman M., Madan A., Young A.C., Shcheglovskoy Y.,
Bouffard G.G., Blakesley R.W., Touchman J.W., Green E.D.,
Dickson M.C., Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
Butterfield Y.S., Krzywinski M.I., Skalska U., Smalios D.E.,
Schnerch A., Schein J.E., Jones S.J. and Marra M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12479932
2 (bases 1 to 2389)
Strausberg R.
Direct Submission
Submitted (11-DEC-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: gcgbs-remail.nih.gov
Tissue Procurement: Gilbert Smith, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Ness, Pawan Pandoh, Anna-Lisa Prabhu, parvaneh Saesedi, Jacqueline
Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Scott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

```

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 131 Row: g Column: 22
 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 27370269.

FEATURES

Location/Qualifiers
 1. 2389
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="Czech II"
 /db_xref="taxon:10090"
 /clone="IWAG:3995710"
 /tissue_type="Mammary tumor metastasized to lung. Tumor arose spontaneously from a senescent normal mammary (clonal) outgrowth infected with the virus MMIV."
 /clone_lib="NCI CGAP Lu29"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 2389;
 Best Local Similarity 100.0%; Pred. No. 5.5e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 |||||
 Db 2339 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2387

RESULT 79
 BC053747
 LOCUS
 DEFINITION Mus musculus solute carrier family 43, member 1, mRNA (cDNA clone MGC:59491 IMAGE:6330614), complete cds.
 ACCESSION BC053747
 VERSION BC053747.1 GI:31753063
 KEYWORDS MGC.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 2507)
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altshuler, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Rana, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullen, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shvachenko, Y., Bouffard, G.G., Blakeley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butlerfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2388457
 MEDLINE
 PUBMED 12477932
 REFERENCE 2 (bases 1 to 2507)
 AUTHORS Strausberg, R.
 TITLE Direct Submission
 JOURNAL Submitted (13-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Mark Maconochie, Ph.D. and Nancy L. Freeman, Ph.D.
 cDNA Library Preparation: ResGen, Invitrogen Corp
 cDNA library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Genome Sequence Centre, BC Cancer Agency, Vancouver, BC, Canada
 info@cgsc.bc.ca
 Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield, Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin, Letticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven Nees, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saeedi, Jacqueline Schein, Duane Smalios, Michael Smith, Lorraine Spence, Jeff Stott, Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy, George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 109 Row: b Column: 18
 This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis.

FEATURES

Location/Qualifiers
 1. 2507
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /clone="MGC:59491 IMAGE:6330614"
 /tissue_type="Embryo, day 9 mouse (CS7BL/6 background) oocytes"
 /clone_lib="NIH MGC_130"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6.1"
 1. 2507
 /gene="Slc43a1"
 /note="synonyms: PB39, R00504"
 /db_xref="LocusID:72401"
 /db_xref="MGI:1931352"
 85. 1779
 /codon_start=1
 /product="Slc43a1 protein"
 /protein_id="AAH83747.1"
 /db_xref="GI:31753064"
 /db_xref="LocusID:72401"

gene

CDS

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 2507;
 Best Local Similarity 100.0%; Pred. No. 5.5e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 |||||
 Db 2324 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2372

RESULT 80

BC053126
 LOCUS
 DEFINITION Danio rerio cDNA clone MGC:63875 IMAGE:6789528, complete cds.
 ACCESSION BC053126

```

VERSION
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS

BC053126.1 GI:31419251
MGC.
Danio rerio (zebrafish)
Danio rerio
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 2582)
Straussberg, R.L., Feingold, B.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altechul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Boutard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalusz, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2582)
Straussberg, R.
Direct Submission
Submitted (02-JUN-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgaps@mail.nih.gov
Tissue Procurement: Leonard I. Zon, M.D.
CDNA Library Preparation: Invitrogen Corp
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www.shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAC Plate: 117 Row: b Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, similarity but not identity to protein.
FEATURES
source
1. .2582
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:63875 IMAGE:6789528"
/cisue_type="Embryo, zebrafish 24 hpf"
/clone_lib="NCI_CGAP_Zemb2"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6.1"
183. .2096
/codon_start=1
/product="Unknown (protein for MGC:63875)"
/protein_id="AAH53126.1"
/db_xref="GI:31419252"
/translation="MNAATAGSYPMASLVGLHDPDITAMLYKFSAGPVLIRVC"

CDS

RDMITRSLGAYVNFQPADAEALDTNFDVVKGKDIRIMWSORDPSLRKSGVGVN
FIKNDKSIDNKALYDTFSAFGNILSKVCDENGSKGYAFVHEFTQDAARAIEKMN
GMLNDRKVFVGRFKSRERAEKAKAEFTNVYIKNFDDMDQRLKELFDKTKGT
LSVKVPTDPTGKSGRGFGVSEYKEDANKAVEENGTGLNGKTVFVGRKQKMERQAE
LKRKVFQKQERISRYQGVNLYIKNLDITIDDEKLKKEKFFSGFTTSKAVMLEERSK
GGFVCFSSPEATKAVTEMGRIYVSKPLVVALAQRKEERKAHLITNOMRIAGMRA
MPANALINQFOPAGGVFVPAQVPAQONRTTYVAPNLTQWRPNRPNWQOQGGQGGFOG
MPSRIRQPGPRANIRHLAPNASQPGMPTGAGVGGTGOPMGRPRMAMGTTTPRAMP
PKYIASTIRNTQFQVVFIAQQAQPAFVHVQOQEBLTASMLAAAPQKQKQMLGERLE
PLIQAMHPSLAGKITGMLLEIDNSLHMLHSHESLRKVEEAVLQHQAKQKATQ
KVGVTAAATAATS"
429. .1301
/misc_feature
/note="COG0724; Region: RNA-binding proteins (REM domain)
[General function prediction only]"
/db_xref="CDD:COG0724"
1809. .2021
/misc_feature
/note="PABP; Region: Poly-adenylate binding protein,
unique domain"
/db_xref="CDD:pfam00658"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 2582;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
Db 2481 AAATGAAAAA 2529

RESULT 81
AX474240 2713 bp DNA linear PAT 12-AUG-2002
LOCUS
DEFINITION
Sequence 1 from Patent EP1223218.
AX474240
ACCESSION
AX474240.1 GI:22213873
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1
AUTHORS
Praser, C.C.
C42000 and cd2001 molecules and uses thereof
Patent: EP 1223218-A 1 17-JUL-2002.
JOURNAL
Millennium Pharmaceuticals, Inc. (US)
FEATURES
Location/Qualifiers
source
1. .2713
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
Db 2650 AAATGAAAAA 2698

RESULT 82
AX474270 2713 bp DNA linear PAT 12-AUG-2002
LOCUS
DEFINITION
Sequence 31 from Patent EP1223218.
AX474270
ACCESSION
AX474270.1 GI:22213882
VERSION
KEYWORDS
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Euthera; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE 1
 AUTHORS Fraser, C.C.
 TITLE Cd2000 and cd2001 molecules and uses thereof
 JOURNAL Patent: EP 1223218-A 31 17-JUL-2002;
 Millennium Pharmaceuticals, Inc. (US)
 FEATURES
 source
 1. .2713
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 5.4e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002
 Db 2650 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002

RESULT 83
 AX474271
 LOCUS 2713 bp DNA linear PAT 12-AUG-2002
 DEFINITION Sequence 32 from Patent EP1223218.
 AX474271
 ACCESSION AX474271.1 GI:22213883
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Fraser, C.C.
 TITLE Cd2000 and cd2001 molecules and uses thereof
 JOURNAL Patent: EP 1223218-A 32 17-JUL-2002;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES

source
 1. .2713
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 5.4e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002
 Db 2650 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002

RESULT 84

AX474272
 LOCUS 2713 bp DNA linear PAT 12-AUG-2002
 DEFINITION Sequence 33 from Patent EP1223218.
 AX474272
 ACCESSION AX474272.1 GI:22213884
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Fraser, C.C.
 TITLE Cd2000 and cd2001 molecules and uses thereof
 JOURNAL Patent: EP 1223218-A 33 17-JUL-2002;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES

source
 1. .2713
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 5.4e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002
 Db 2650 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002

RESULT 85

AX474273
 LOCUS 2713 bp DNA linear PAT 12-AUG-2002
 DEFINITION Sequence 34 from Patent EP1223218.
 AX474273
 ACCESSION AX474273.1 GI:22213885
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Fraser, C.C.
 TITLE Cd2000 and cd2001 molecules and uses thereof
 JOURNAL Patent: EP 1223218-A 34 17-JUL-2002;
 Millennium Pharmaceuticals, Inc. (US)

FEATURES

source
 1. .2713
 /organism="Homo sapiens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:9606"

ORIGIN

Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 5.4e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002
 Db 2650 AAATGAAAAA 2713 bp DNA linear PAT 12-AUG-2002

RESULT 86

BC053348
 LOCUS 2789 bp mRNA linear PRI 08-OCT-2003
 DEFINITION Homo sapiens solute carrier family 22 (organic anion/cation transporter), member 12, transcript variant 1, mRNA (cdna clone MGC:61679 IMAGE:5183650), complete cds.
 BC053348
 ACCESSION BC053348.1 GI:31419813
 VERSION
 KEYWORDS
 SOURCE MGC.
 ORGANISM Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G., Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D., Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Shat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheer, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,

Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krawinski, M.I., Skaleka, U., Smalil, D.E., Schnerch, A., Schain, J.E., Jones, S.J., and Marra, M.A.,
Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

2338257
12477932
2 (bases 1 to 2789)
Straussberg, R.
Direct Submission
Submitted (02-JUN-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
Contact: MGC help desk
Email: cgabs@mail.nih.gov
Tissue Procurement: Life Technologies, Inc.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland
Web site: <http://www.nisc.nih.gov/>
Contact: nisc.mgc@nih.gov
Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, O.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Teague, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 115 Row: c Column: 21
This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 24497484.

Location/Qualifiers
1. .2789
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:61679 IMAGE:5183650"
/tissue_type="Colon, Kidney, Stomach, adult, whole pooled"
/clone_lib="NIH MGC 116"
/lab_host="DH10B"
/notes="Vector: pCMV-SPORT6"
1. .2789
/genes="SLC22A12"
/notes="synonyms: OAT4L, URAT1, RST"
/db_xref="LocusID:116085"
/db_xref="MIM:607096"
291. .1952
/codon_start=1
/product="urate anion exchanger 1, isoform a"
/protein_id="AAH53348.1"
/db_xref="GI:31419814"
/db_xref="LocusID:116085"
/translations="MAFSLDLVGLGRFOVLQTMALMVSIMVLTOSMLNFSAAV
PSRCHAPLDNSTAGSLGSLSPALIAISIPFGNORPHCRFRFPQWQLDPN
ATATSEADTEPCVDGWYDRSIFTSTIVAKNVLCDSHLKPQAQSIYLAGILVGA
AACGDSDFRGLVLTWGYLQVAVNGTAAPAPFPVYCLFRFLFAVAGVMNTG
TLMETAAARPLVMTSLGSLGFGHGLTAAGVAVGVDRVTLQLVSVFPFLCPLYS
WMLAESARLLTGRLDWGLQELRVAALNGKGVODTLTPVLLSAMREELSMGQPP
ASGLTLRVRGPRCTISTLCWFAFGFTFGGLDLOALGNSIFLQWFGVVDIPA
KMGALLLHSLGRPTLARSLLAGLCILANTLVHEMGALASALVILGVGGAFT
CITYISELFTVLRMTAVGLGMAAGGAILGFLVGLVGVGFWPLLVITVTPVLS
GLAALLPFTQSLPDLPTQDVQVQAKKATHGTGNSVLRKSTQF"
/note="sugar_tr; Region: Sugar (and other) transporter"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 2803;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

misc_feature
574. .575
/note="splice junction"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 2803;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 2789;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762
|||||
Db 2739 AAATGAAA 2787

RESULT 87
GDFIMB LOCUS Chicken mRNA for fimbrin. 2803 bp mRNA linear VRT 12-SEP-1993
DEFINITION X52562
ACCESSION X52562.1 GI:62889
VERSION
KEYWORDS fimbrin.
SOURCE Gallus gallus (chicken)
ORGANISM Gallus gallus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Archosauria; Aves; Neognathae; Galliformes; Phasianidae;
Phasianinae; Gallus.
REFERENCE 1 (bases 1 to 2803)
de Arruda, M.V., Watson, S., Lin, C.S., Leavitt, J. and Matsudaira, P.
Fimbrin is a homologue of the cytoplasmic phosphoprotein plastin
and has domains homologous with calmodulin and actin gelation
proteins
J. Cell Biol. 111 (3), 1069-1079 (1990)
JOURNAL 90361735
MEDLINE 2391360
PUBMED
REFERENCE 2 (bases 1 to 2803)
de Arruda, M.V.
AUTHORS
TITLE Direct Submission
JOURNAL
FEATURES
source
1. .2803
/organism="Gallus gallus"
/mol_type="mRNA"
/db_xref="taxon:9031"
/clone="M38,M48,SW6"
/cell_type="epithelial"
/tissue_type="intestine"
78. .1970
/notes="unnamed protein product; fimbrin (AA 1-630)"
/codon_start=1
/protein_id="CAA36796.1"
/db_xref="GI:62890"
/db_xref="GOA:P19179"
/db_xref="SWISS-PROT:P19179"
/translations="MENNVTTSRELEELREAFNKIDINSQVSDYELQDLFKEAS
LPLPGYKVRILEIKFPAVDSNDKGNPEEFVSLIOELKSDVSKYRSINKKIGI
TALGCTSSISTEGTQHSYSEEEKVAFVWINKALQDDPCKHLPNPSDASLFXSLA
QEGKPLVLGLWQIIKVGFLADIIRNEALJALLNEGEELQMKLSPEELLRWV
NYHLNAGWKIISFQSDIRDSRAYYHLLNQIAFGDDEIHEVDFSGFNDDLR
RAECMLQADKLCRCQFVTPADVAGNPKLNLAVANLFTYFALHKKPDNSYDLTLL
EGSNERTFRNWNLSGVSPYVNHVSLSDALIIFOLYEMTRVPVDTWVTKRPP
LLGNMKKISCNVAVELGKTSFVGLAGHDLNEGPTLTLALIQLMRYTLNV
LSDLGEGKYNDRIIIKWNQTLANANKTSITSFKDKSISTSLPVLDLDAIPKAV
RQENKREDLSYQDKLNKAKIASVARKIARIALPDDLVLEKPMVTVFACLMGR
GLANKIK"
574. .575
/note="splice junction"

```

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3761
Db 2755 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2803

RESULT 88
LOCUS BC014041 Homo sapiens KIAA1404 protein, mRNA (CDNA clone IMAGE:4546448), 2810 bp mRNA linear PRI 05-NOV-2003
DEFINITION partial cds.
ACCESSION BC014041
VERSION BC014041.2 GI:33874124
KEYWORDS Homo sapiens (human)
SOURCE Homo sapiens
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 2810)
Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.W., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Kzyvinski,M.I., Skalska,U., Smalish,D.E.,
Schmerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 2810)
Strausberg,R.
Direct Submission
Submitted (10-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:15559352.
Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),
Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc\_mgc@nri.nih.gov
Ahter,N., Ayele,K., Beckstrom-Sternberg,S.M., Benjamin,B.,
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dierich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-L., Karlins,E., Kwong,P., Laric,P., Legaspi,R.,
Maduro,Q.L., Mastillo,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantripop,S., Thomas,P.J., Touchman,J.W.,
Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRAL Plate: 26 Row: p Column: 19.
Location/Qualifiers

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 2810;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 2750 AATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2798

RESULT 89
LOCUS AR216297 2851 bp DNA linear PAT 25-SEP-2002
DEFINITION Sequence 1 from patent US 6410714.
ACCESSION AR216297
VERSION AR216297.1 GI:23314767
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 2851)
AUTHORS Weber,E.R. and McCall,C.A.
TITLE Canine low affinity Ige receptor (CD23) nucleic acid molecules and
uses thereof
JOURNAL Patent: US 6410714-A 1 25-JUN-2002;
FEATURES Location/Qualifiers
source 1. 2851
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

1. 2810
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4546448"
/tissue_type="Colon, adenocarcinoma"
/clone_lib="NIH MGC 15"
/lab_host="DH10B-R"
/notes="Vector: pOTB7"
1. 2810
/gene="KIAA1404"
/notes="Synonym: FLJ11277"
/db_xref="LocusID:57169"
64. 2810
/gene="KIAA1404"
/codon_start=1
/product="KIAA1404 protein"
/protein_id="AAH1404.1"
/db_xref="GI:33874125"
/db_xref="LocusID:57169"
/translation="MEERRPHLDARPNRSHNTHRGVDELPPLPPARNQANPPANALR
GGASHRGPRANHPAAAYWOREEPRFRAMGRNPHOGRNQHGHASDEARDORHDOEND
TGWRRGNODCRNRPWNSDNFQOWRTPHQKTEPOQAKLGYKFLSLQKDFSEV
VITLNSGLKELLSHSHKSNFLLELCQVLRKAGSSKMDRQSVLHVILGNKSLFK
VCLPAYVGMITEPIPDINQVPEHISNLSLLQDLNVSFFPASSVQETSMLVSLPTS
LNALRASVDIETEKNLEKVQTHIEHQKREGTLKRVDTYTLVQPEAEHUVESYR
TMPIYPTNVEHLDRPFLRNIIISKYDSTAIIYDTHFLRLREDFVRLREGILLEL
QSFEQGLRKRFDDIRIVFDTRIITPMCSSGIYVQFDTKPLKFWQNSKELLY
SLVCMKDNFETFLPATVSNREQDLKRGIVQLCFNEQSQQLAEVQSPDSFLNPSA
TAYFAYRHVLEGLQVQVEEDVPFORNIVECNHSHVKEPYLLMGGRYDFTPLIENFA
TGEFLRNVGLRHPRINVLDPQOWSKALKLDSDSOMALQFALRELAIQGPPTG
KTYGKIVQALLTNSVWSISQKFPFIVVYCVYTHALDQFLEGIYNQKTSIVAVGG
RNSGLIKOFTLRELNRKFRPNMHLRRAYNSINTOMKESQELQEGHGAKILSCTM
RGVLRQYLYKISQHWESLNGPQVSEWTCFQHWKSHMMLWGLVGLGSGTFSVS
PAGPENTAQEGDEEEEGEESLIEIAEADLIQADRVIEEVEVVRPQRKKESGA
DOELAKMLLAMELDHCGTGTAAQEQATCEMOTQRNKKKKKKKKKKKKKKKK
K"

```

```

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 2802 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2850

RESULT 90
LOCUS AR216298 2851 bp DNA linear PAT 25-SEP-2002
DEFINITION Sequence 3 from patent US 6410714.
ACCESSION AR216298
VERSION AR216298.1 GI:23314768
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
AUTHORS Weber, E.R. and McCall, C.A.
TITLE Canine low affinity IGE receptor (CD23) nucleic acid molecules and
uses thereof
JOURNAL Patent: US 6410714-A 3 25-JUN-2002;
FEATURES
source Location/Qualifiers
1..2851
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN
Query Match 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 50 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2

RESULT 91
LOCUS BC060912 2954 bp mRNA linear VRT 04-NOV-2003
DEFINITION Danio rerio cDNA clone IMAGE:4144592, partial cds.
ACCESSION BC060912
VERSION BC060912.1 GI:38173866
KEYWORDS
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
1 (bases 1 to 2954)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Udén, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullahy, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Mazny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzyzanski, M.I., Skalska, U., Smalilus, D.E.,
Scherch, A., Schein, J.E., Jones, S.J. and Matra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
REFERENCE
AUTHORS Straussberg, R.
TITLE

Direct Submission
Submitted (03-NOV-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs@mail.nih.gov
Tissue Procurement: Susan E. Brockerhoff, University of Washington
cDNA Library Preparation: Susan E. Brockerhoff, University of
Washington
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAK Plate: 136 Row: b Column: 8.
FEATURES
source Location/Qualifiers
1..2954
/organism="Danio rerio"
/mol_type="mRNA"
/strains="wild-type"
/db_xref="taxon:7955"
/clone="IMAGE:4144592"
/tissue_type="Eye, adult retina, 1-2 year old, mixed sex"
/clone_lib="Zebrafish adult retina cDNA"
/lab_host="DH10B"
/notes="Vector: pBluescript SK-"
<1..1333
/codon_start=2
/product="Unknown (protein for IMAGE:4144592)"
/protein_id="AAH60912.1"
/db_xref="GI:38173867"
/translacion="SLLPSFCFOLDSEBLESKHTSTWEIDTLGAFDRFEKKGEKES
CPILDFLHVAKTGTMTIQWQKSYFLKLEKVMDDFKASCPQGRGPANPNVYIP
PEMKQRILKIVNGYNGIPTIQLCELLTEPKRYNAGTDKFLRGVKNVMVVCVYP
TESKQSCVNRNGVVFENISAFDRNVNGPGTPELRNRPKHSLSNVANGLPDS
TESKQASQSESTVNESASAEASHKRRDDSDATHEAKRLEKFEKKEE
EEDDEDEDEDEIKELDEPHSCTSVAESSDVPQSSTDTVAEKDDEDEPSTQ
SEVENVGRSTSESDSPDPSHKATGESDPKQAESESESAEQTEETNDPVSS
SSNNSDEGVSSAETPSAPSSSTELPAEGSVTABITSDNSETADDNMEQD"

ORIGIN
Query Match 1.3%; Score 49; DB 5; Length 2954;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 2905 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2953

RESULT 92
LOCUS BC029643 3049 bp mRNA linear ROD 11-DEC-2003
DEFINITION Mus musculus RIKEN cDNA 2810408E11 gene, mRNA (cDNA clone MGC:25768
IMAGE:4010810), complete cds.
ACCESSION BC029643
VERSION BC029643.1 GI:20987279
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 3049)
Straussberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,

```

Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Scapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Fehny, J., Heiton, E., Kettelman, M., Madan, A., Rodrigues, S., Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.E., Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

12477932

2 (bases 1 to 3049)

Direct Submission

Submitted (06-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILML)

DNA Sequencing by: Baylor College of Medicine Human Genome Sequencing Center

Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>

Contact: amg@bcm.tmc.edu

Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H., Kowis, C.R., Sneed, A.U., Martin, R.G., Muzny, D.M., Nanavati, A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILML at: <http://image.llnl.gov>

Series: IPAK Plate: 30 Row: i Column: 17

This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORF analysis, Similarity but not identity to protein.

FEATURES

source

1..3049

/organism="Mus musculus"

/mol_type="mRNA"

/strain="Czech II"

/db_xref="taxon:10090"

/clone="MGC:25768 IMAGE:4010810"

/tissue type="Mammary tumor metastatized to lung."

MMTV-LTR/Mnt1 model. Expression driven by an MMTV-LTR enhancer."

/clone_lib="NCI CGAP_Lu30"

/lab_host="DH10B"

/note="Vector: PCMV-SPT6"

1..3049

/gene="2810408E11Rik"

/db_xref="LocusID:106200"

/db_xref="MGI:1923620"

24..2870

/codon_start=1

/product="expressed sequence A1427833"

/protein_id="AAH29643.1"

/db_xref="GI:20987280"

/db_xref="LocusID:106200"

/translation="MSECGRGGGGNNSEDAEBGGPGKSGSLSPAGAAASLEGRIRGLGASIMAFQRPELLCGAVALGKFTCSRAKDVIPAKPPVFFSRRSPVLDLFGQLDYADHVRQSEVVLFYAPWCGQSIARAEIEQAASRLSDQVLVAIN

gene

CDS

CMWNGKCRKQKHFFYPVHLYHRSFGPIETKYGPMASVYIEKFRVRAKMLLYIPSQ
SALUDFSSVEPGVLGYEFPSPGPPGYLTFTFSAHLSKDKDYLGTVRFGVITDKHL
SALVLSVSGSVLYHREFTSLVFPREVMFTAEINIKMASENQETLFRVLOPHGGKS
LLLNNEKGPALFLFPDLARHLLDEITVEALEVNNCHGDOVVERLLOHREV
EAPVLSLAPLSPDTQMLASPCCNTIVILPQGPALSRTHNVCLCYNQTVGGTR
PSSVSPQCFEFMAALDSFYLBKEQTFYHVSIGSIECSNPLTISFSTYIACRRI
SRGASFTGEQNVLTAPAEFSELEKCEATAPSSPIEENRYRFPQGLTSTAF
GLSCTRNTLNIIYLDNSLFWLAERLGAAPSAPVKFPAFIVDKESHVILDPQAL
MYKTFESFIONFVLYSLKRLHLSAQFPTOHLITEVTTDTFWEVTLRKQDVLL
VYTWQCFPSLNHIFQLARLLPDTFVARIDVSONDLFPWFMDRLPVLFPFN
KDLVXPYGDLPITLNLARLILHLSDAASAPDGPISPTQCVOSKVALOREHLS
HVENAMKREMSLSLARTQVEGRLLSARDGRLLRQRTLEQGRLLRHSOKL
QALYIKKARELOELASGTPPEHTWLKILVATMERELECGGAKAPLKGKAPNH
SKSGTGTLPDGTPTTSTSLASTKHENRTD"

misc_feature

357..644

/note="Thioredoxin; Region: Thioredoxin. Thioredoxins are small enzymes that participate in redox reactions, via the reversible oxidation of an active centre disulfide bond. Some members with only the active site are not separated from the noise"

/db_xref="CDD:pfam00085"

2007..2327

misc_feature

/note="Thioredoxin; Region: Thioredoxin. Thioredoxins are small enzymes that participate in redox reactions, via the reversible oxidation of an active centre disulfide bond. Some members with only the active site are not separated from the noise"

/db_xref="CDD:pfam00085"

ORIGIN

Query Match 1.3%; Score 49; DB 10; Length 3049;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

Db 2971 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3019

RESULT 93

AB063100 3072 bp mRNA linear PRI 13-JUN-2001
LOCUS Macaca fascicularis brain cDNA clone:Qt-A-14469, full insert
DEFINITION sequence.

ACCESSION AB063100

VERSION AB063100.1 GI:14388574

KEYWORDS oligo capping; fis (full insert sequence).

SOURCE Macaca fascicularis (Crab-eating macaque).

ORGANISM Macaca fascicularis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea; Cercopithecinae; Macaca.

REFERENCE 1 (sites)

Osada, N., Hida, M., Kusuda, J., Tanuma, R., Iseki, K., Hirai, M.,

Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.

Isolation of full-length cDNA clones from macaque brain CDNA

libraries

Unpublished

2 (bases 1 to 3072)

Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.

Direct Submission

Submitted (11-JUN-2001) Kateyuki Hashimoto, National Institute of

Infectious Diseases, Division of Genetic Resources, 23-1, Toyama

1-chome, Shinjuku-ku, Tokyo 162-8640, Japan

(E-mail: khashi@nih.go.jp, URL: <http://www.nih.go.jp/yoken/genebank/>,

Tel: 81-3-5285-1111 (ex.2120), Fax: 81-3-5285-1181)

Lab host: TOP10

Vector: pME18S-Fl3 (Acc. No. AB009864)

R. Site1: Drallii (CAGCTGTGTG)

R. Site2: Drallii (CAGCTGTGTG)

Description: 1st strand cDNA was primed with an oligo(dT) primer [ATGTGGCTTTTTTTTTTTT]; double-stranded cDNA was synthesized using specific 5' and 3' primers and amplified by PCR. The PCR

product was digested with SfiI and size selection was performed to exclude fragments <1.5kb. The SfiI-digested PCR product was cloned into distinct DraIII sites of pME18S-FL3. XhoI sites just outside the DraIII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., Institute of Medical Science, University of Tokyo).

Custom primer used for sequencing
(5' end primer [CTTCTGCTCTAAAGTGGG];
3' end primer [CGACTGCGAGTCGAGCACA]).

FEATURES

source

```
1..3072
  /organism="Macaca fascicularis"
  /mol_type="mRNA"
  /db_xref="taxon:9541"
  /clone="Qtra-14469"
  /sex="male"
  /tissue_type="temporal lobe right"
  /clone_lib="macaque brain cDNA library Qtra"
  /dev_stages="adult"
267..1211
  /codon_start=1
  /product="hypothetical protein"
  /protein_id="BA860806.1"
  /db_xref="GI:14398575"
  /translation="MGLDSEFGSVLVNVSTALNDTFEYRTWTSIADKRVENWPLMQ
SPWPTLSISTLYLLEFVWVGPVKMDREFQMLVLIYFNGVLLNFFIFPELPMGSI
NAGYSYIQSDVSYNNVNEVIRAAALWYFVSKGYELDTVFFILRKNQVSLHYV
HCTMTLWIGIKVAGGAGFFGAGQNSFIVIMSYGLAAGPMTIKYLMWKRVL
TMQLQVQHVHTIGHTSLYTDCTPPKMWHALIAYALSIFLFLNFIYRTYKPKXP
KTGTAMNGISANGVSKSEKQLVIENGKKQXGKAKGD"
```

ORIGIN

```
Query Match      1.3%; Score 49; DB 9; Length 3072;
Best Local Similarity 100.0%; Pred. No. 5.4e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 2991 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3039
```

RESULT 94

```
BC054927
LOCUS          3124 bp mRNA linear VRT 07-OCT-2003
BC054927       Danio rerio cDNA clone MGC:63655 IMAGE:5612000, complete cds.
ACCESSION     BC054927.1 GI:32766369
VERSION       MGC.
SOURCE        Danio rerio (zebrafish)
ORGANISM      Danio rerio
               Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
               Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
               Cypriniformes; Cyprinidae; Danio.
REFERENCE     1 (bases 1 to 3124)
               Strausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
               Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
               Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
               Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
               Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
               Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
               Schetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
               Carinci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
               Abramson,R.D., Mullahy,S.J., Bosak,S.A., McEwan,P.J.,
               McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
               Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
               Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
               Fahy,J., Helton,E., Kettner,M., Madan,A., Rodrigues,S.,
               Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
               Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
               Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
               Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
               Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
```

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3124)
Strausberg,R.
Direct Submission
Submitted (08-JUL-2003) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK

COMMENT

Contact: MGC help desk
Email: cgabbs@mail.nih.gov
Tissue Procurement: Dr. Sumio Sugano
cDNA Library Preparation: Dr. Sumio Sugano
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@bcsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutache, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeay, Steven
Ness, Pawan Pandori, Anna-Liisa Prabhoo, Parvaneh Saeedi, Jacqueline
Schein, Duane Smalilus, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Minada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAC Plate: 118 Row: 0 Column: 1
This clone was selected for full length sequencing because it passed the following selection criteria: Hexamer frequency ORP analysis, Similarity but not identity to protein.

FEATURES

source

```
1..3124
  /organism="Danio rerio"
  /mol_type="mRNA"
  /db_xref="taxon:7955"
  /clone="MGC:63655 IMAGE:5612000"
  /tissue_type="whole body, adult male"
  /clone_lib="Sugano SJD adult male"
  /lab_host="DH10B"
  /note="Vector: pME18S-FL3"
222..890
  /codon_start=1
  /product="Unknown (protein for MGC:63655)"
  /protein_id="AAH54927.1"
  /db_xref="GI:32766370"
  /translation="MDRISRYLSRVYGMHPKDTARQLVLAQDGLVLTITVCGKSGK
AGTEQGYWLPQDEMGSEKPSLPSEALFPDDWEASHDWYCFCHLPQGVMECD
GCFRVYHLRCVSEDRPRTTSHWCGICRGSKRKNLKNQENTVYKFLILGRKRAV
DLKSFEEFRADQLIVHTAILNGVNSQTEIARLLYNDTCLLASCGFVFCARSPSD
WGPA"
misc_feature     483..617
  /note="PHD; Region: PHD zinc finger"
  /db_xref="CDD:smart00249"
misc_feature     627..818
  /note="BROMO; Region: bromo domain"
  /db_xref="CDD:smart00297"
```

ORIGIN

```
Query Match      1.3%; Score 49; DB 5; Length 3124;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 3023 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3071
```

```

RESULT 95
BC036825      3145 bp      mRNA      linear      PRI 26-AUG-2002
LOCUS      Homo sapiens, clone IMAGE:5191998, mRNA, partial cds.
DEFINITION
ACCESSION      BC036825
VERSION      BC036825.1  GI:22477362
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Strausberg, R.
TITLE      Direct Submission
JOURNAL      Submitted (23-AUG-2002) National Institutes of Health, Mammalian
              Gene Collection (MGC), Cancer Genomics Office, National Cancer
              Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
              USA
REMARK
COMMENT      NIH-MGC Project URL: http://mgc.nci.nih.gov
              Contact: MGC help desk
              Email: cgsbbs-remail.nih.gov
              Tissue Procurement: Life Technologies, Inc.
              cDNA Library Preparation: Life Technologies, Inc.
              cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
              DNA Sequencing by: National Institutes of Health Intramural
              Sequencing Center (NISC),
              Gaithersburg, Maryland;
              Web site: http://www.nisc.nih.gov/
              Contact: nisc.mgc@nih.gov
              Akhter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
              Blakeley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
              Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
              Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Lalic, P., Legaspi, R.,
              Maduro, O.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
              McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
              Tsurgoon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
              Young, A., Zhang, L.-H. and Green, E.D.
              Clone distribution: MGC clone distribution information can be found
              through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
              Series: IRAK Plate: 78 Row: e Column: 23
              This clone was selected for full length sequencing because it
              passed the following selection criteria: Hexamer frequency ORF
              analysis.
FEATURES
source
1..3145
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5191998"
    /tissue_type="Brain, adult, 6 pooled whole brains"
    /clone_lib="NIH MGC_114"
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
    /c1..692
    /codon_start=3
    /product="Unknown (protein for IMAGE:5191998)"
    /protein_id="AAH36825.1"
    /db_xref="GI:22477363"
    /translation="CLQSHSQPSQLLCTLLCWLGRERPRFKAQSTSPGLEGVPRF
    LKRLLLTWGCVSVWAVQVDIFTAGKLLALALIIIMGIQVICKGYFVLEPKNAFE
    NQFQEDLVGVALAFQGSFAYGWNFLNVTEELVDVFKLFFAGAREGHLPSPVLAMTH
    VKRCTPIPALFTCTILLMLVTSMTYLLINVGVFINVLYFVGVTVAGQIVLRWKKPDI
    PRPIKPLFIQG"
CDS
1..3145
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="IMAGE:5191998"
    /tissue_type="Brain, adult, 6 pooled whole brains"
    /clone_lib="NIH MGC_114"
    /lab_host="DH10B"
    /note="Vector: pCMV-SPORT6"
    /c1..692
    /codon_start=3
    /product="Unknown (protein for IMAGE:5191998)"
    /protein_id="AAH36825.1"
    /db_xref="GI:22477363"
    /translation="CLQSHSQPSQLLCTLLCWLGRERPRFKAQSTSPGLEGVPRF
    LKRLLLTWGCVSVWAVQVDIFTAGKLLALALIIIMGIQVICKGYFVLEPKNAFE
    NQFQEDLVGVALAFQGSFAYGWNFLNVTEELVDVFKLFFAGAREGHLPSPVLAMTH
    VKRCTPIPALFTCTILLMLVTSMTYLLINVGVFINVLYFVGVTVAGQIVLRWKKPDI
    PRPIKPLFIQG"
ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 3145;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      3714 AAATGAAAAA...
DB      3060 AAATGAAAAA...

```

```

RESULT 96
HSM800211      3174 bp      mRNA      linear      PRI 18-FEB-2000
LOCUS      Homo sapiens mRNA; cDNA DKFZp586C241 (from clone DKFZp586C241).
DEFINITION
ACCESSION      AL049426
VERSION      AL049426.1  GI:4500206
KEYWORDS
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS      Wambutt, R., Heubner, D., Mewes, H.W., Gassenhuber, J. and Wiemann, S.
TITLE      Direct Submission
JOURNAL      Submitted (10-MAR-1999) MIPS, Am Klopferspitz 18a, D-82152
              Martinsried, GERMANY
COMMENT      Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
              Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
              sequenced by AGOWA (Berlin/Germany) within the cDNA sequencing
              consortium of the German Genome Project.
              This clone (DKFZp586C241) is available at the RZPD in Berlin.
              Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
              Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de Further
              information about the clone and the sequencing project is available
              at http://www.mips.biochem.mpg.de/proj/cDNA/.
FEATURES
Location/Qualifiers
1..3174
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="RZPD:DKFZp586C241"
    /db_xref="taxon:9606"
    /clone="DKFZp586C241"
    /tissue_type="uterus"
    /clone_lib="586 (synonym: hute1). Vector pSport1; host
    DH10B; sites NotI + SalI/MluI"
    /dev_stage="adult"
    /dev_Stage="3104..3109"
    /polyA_signal
    /polyA_site
ORIGIN
Query Match      1.3%; Score 49; DB 9; Length 3174;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY      3714 AAATGAAAAA...
DB      3119 AAATGAAAAA...

```

```

RESULT 97
BC032176      3290 bp      mRNA      linear      ROD 16-APR-2003
LOCUS      Mus musculus RIKEN CDNA 953006B08 gene, mRNA (cDNA clone
DEFINITION      IMAGE:5322259), partial cds.
ACCESSION      BC032176
VERSION      BC032176.1  GI:21595134
KEYWORDS
SOURCE      Mus musculus (house mouse)
ORGANISM      Mus musculus
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
AUTHORS      Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
              Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
              Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
              Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
              Diatchenko, L., Marusik, K., Farmer, A.A., Rubin, G.M., Hong, L.,
              Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
              Scheetz, T.E., Brownstein, M.J., Ustin, T.B., Toshiyuki, S.,
              Carninci, P., Prange, C., Raha, S., Loquellano, N.A., Peters, G.J.,
              Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

```

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalish, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 2388257
 12477932
 2 (bases 1 to 3290)
 Strausberg, R.
 Direct Submission
 Submitted (06-JUN-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@hgrl.nih.gov
 Akter, N., Ayele, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granate, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Latic, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsugeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 56 Row: 1 Column: 17
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES
 source
 1. .3290
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5322259"
 /tissue_type="Mammary tumor, C3(1)-Tag model. Infiltrating
 ductal carcinoma. 5 month old virgin mouse."
 /clone_lib="NCI_CGAP_Name"
 /lab_host="DH10B"
 /note="Vector: pCMV-SPORT6"
 <1. .3290
 /gene="9530006B08Rik"
 /db_xref="LocusID:229584"
 /db_xref="MGI:2442117"
 <1. .2696
 /gene="9530006B08Rik"
 /codon_start=3
 /product="9530006B08Rik protein"
 /protein_id="AAH32176.1"
 /db_xref="GI:21595135"
 /db_xref="LocusID:229584"
 /db_xref="MGI:2442117"
 /translation="TRPLDQNGEVDGHTICOHYVRQSTFFQCHLENVHVSPEST
 TKCKICENAFSEPLFQHKMDTHKGPMPYVCQVRSLSYSEVDVHFRMIHEDTR

ORIGIN
 Query Match 1.3%; Score 49; DB 10; Length 3290;
 Best Local Similarity 100.0%; Pred. No. 5.3e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3714 AAATGAAA 3762
 DB 3242 AAATGAAA 3290
 RESULT 98
 BC034128
 LOCUS
 DEFINITION
 Mus musculus, Similar to neuronal specific transcription factor
 DAT1, clone IMAGE:4913908, mRNA, partial cds.
 ACCESSION
 BC034128
 VERSION
 BC034128.1 GI:21706632
 KEYWORDS
 Mus musculus (house mouse)
 Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus;
 Strausberg, R.
 Direct Submission
 Submitted (02-JUL-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgabs-remail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) mcd@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.
 Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 44 Row: 1 Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Hexamer frequency ORF
 analysis.

FEATURES
 source
 1. .3387
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4913908"
 /tissue_type="Salivary gland, 10 week old female mouse"
 /clone_lib="NCI_CGAP_SG2"
 /lab_host="DH10B"

```

/note="Vector: pQWV-SPORT6"
<1..341
/codon_start=3
/product="Similar to neuronal specific transcription
factor DAIL"
/protein_id="AAH34128.1"
/db_xref="GI:21706633"
/translation="HAGSLKACCCDRLGVSTLYTKANLILCRDRYLRLFGVTCNC
AAKSLIPAPEMVRKDNVYHLDFACQLCNORFCVGDFFLKNMILCQTDYBGL
MKEGYAQRV"

ORIGIN
Query Match 1.3%; Score 49; DB 10; Length 3387;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA... 3563 bp mRNA linear PRI 03-OCT-2003
Homo sapiens syndecan 3 (N-syndecan), mRNA (cdna clone MGC:12759
IMAGE:4121937), complete cds.
BC013974
ACCESSION BC013974
VERSION 1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 3563)
Strausberg,R.D., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Haieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Schetz,T.B., Brownstein,M.J., Usdin,T.B., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahy,J., Helton,E., Kettner,M., Madan,A., Rodriguez,S.,
Sanchez,A., Whitting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Boutard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,Y.S., Krzywinski,M.I., Skalska,U., Smalilus,D.E.,
Scherch,A., Schein,J.E., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
2 (bases 1 to 3563)
Strausberg,R.
Direct Submission
Submitted (10-SEP-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgi.nci.nih.gov
On Aug 19, 2003 this sequence version replaced gi:15559238.
Contact: MGC help desk
Email: gcapbs@mail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: Rubin Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: National Institutes of Health Intramural
Sequencing Center (NISC),

```

```

Gaithersburg, Maryland;
Web site: http://www.nisc.nih.gov/
Contact: nisc_mgc@nri.nih.gov
Blakesley,R.W., Bouffard,G.G., Breen,K., Brinkley,C., Brooks,S.,
Dietrich,N.L., Granite,S., Guan,X., Gupta,J., Haghighi,P.,
Hansen,N., Ho,S.-I., Karlins,E., Kwong,P., Latic,P., Legaspi,R.,
Maduro,Q.L., Masello,C., Maskeri,B., Mastrian,S.D., McCloskey,J.C.,
McDowell,J., Pearson,R., Stantrippop,S., Thomas,P.J., Touchman,J.W.,
Tsurgon,C., Vogt,J.L., Walker,M.A., Wetherby,K.D., Wiggins,L.,
Young,A., Zhang,L.-H. and Green,E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAU Place: 17 Row: 1 Column: 19
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 7662137.

FEATURES
Location/Qualifiers
1..3563
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:12759 IMAGE:4121937"
/issue_type="Muscle, rhabdomyosarcoma"
/clone_lib="NIH MGC 17"
/lab_host="DH10B-R"
/note="Vector: pQWV"
1..3563
/gene="SDC3"
/note="synonyms: KIAA0468, SDCN, SYND3"
/db_xref="LocusID:9672"
325..1437
/codon_start=1
/product="SDC3 protein"
/protein_id="AAH13974.1"
/db_xref="GI:15559239"
/db_xref="LocusID:9672"
/translation="NAIAYIGSCSPSSLPSSLSPTTDFEESGIETAMRSPD
VALAVSTTFEELPSERPLEATPLVVTVEPSPQATVTTMTATATSTGDP
TVAATVATATATPTTAAAPTATTATVTTGVRLLPLTTTATATATATPEASP
PTTAAVLTDEATPTPLVATATSPALPRATPTQEDIPERSTPLGTATPGTTEVAQ
TPFPEFLTITRDEPEVPSGPGDFELPEBETPTQPTANVAVGVGAAGKASPPG
TLPGARPGGLDNDADSGSAQQLPKSILIERKVLVAVIVGVGVGALFAAFVTL
LIYRMKKDEGYSYIEEPQASVYQKPKQKEFYA"
826..1428
/note="Syndecan; Region: Syndecan domain. Syndecans are
transmembrane heparin sulfate proteoglycans which are
implicated in the binding of extracellular matrix
components and growth factors"
/db_xref="CDD:pfam01034"

ORIGIN
Query Match 1.3%; Score 49; DB 9; Length 3563;
Best Local Similarity 100.0%; Pred. No. 5.3e-13;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA... 3608 bp mRNA linear ROD 16-APR-2003
Db 3495 AAATGAAAAA... 3543

RESULT 100
BC028880
LOCUS BC028880
DEFINITION Mus musculus apolipoprotein B, mRNA (CDNA clone IMAGE:4163231),
partial cds.
ACCESSION BC028880
VERSION BC028880.1 GI:20810021
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 3608)
 Search completed: April 23, 2004, 08:22:46
 Job time : 14172 secs

REFERENCE
 AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
 Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
 Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
 Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
 Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
 Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
 Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
 Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
 Abramson, R.D., Mullaly, S.J., Bosak, S.A., McEwan, P.J.,
 McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
 Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
 Villalón, D.K., Muzny, D.N., Sodergren, E.J., Lu, X., Gibbs, R.A.,
 Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
 Sanchez, A., Whitting, M., Madan, A., Young, A.C., Shevchenko, Y.,
 Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
 Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
 Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalusz, D.E.,
 Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
 Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences
 Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
 22388257
 12477932
 2 (bases 1 to 3608)
 Strausberg, R.
 Direct Submission
 Submitted (01-MAY-2002) National Institutes of Health, Mammalian
 Gene Collection (MGC), Cancer Genomics Office, National Cancer
 Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
 USA
 NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 Contact: MGC help desk
 Email: cgapps-r@mail.nih.gov
 Tissue Procurement: Jeffrey E. Green, M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Baylor College of Medicine Human Genome
 Sequencing Center
 Center code: BCM-HGSC
 Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: amg@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Loulseged, H.,
 Kowis, C.R., Sneed, A.J., Martin, R.G., Muzny, D.N., Nanavati,
 A.N., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAC Plate: 37 Row: d Column: 20
 This clone was selected for full length sequencing because it
 passed the following selection criteria: Similarity but not
 identity to protein.

FEATURES
 Source
 1..3608
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:4163231"
 /tissue_type="Liver, normal. 5 month old male mouse."
 /clone_lib="NCI CGAP_Li9"
 /lab_host="DH108"
 /note="Vector: PCMV-SPORT6"

ORIGIN

Query Match 1.3% Score 49 DB 10 Length 3608;
 Best Local Similarity 100.0% Pred. No. 5.3e-13;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762
 |||||||
 Db 3548 AAATGAAA 3596

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 03:54:40 ; Search time 244 Seconds
(without alignments)
8556.251 Million cell updates/sec

Title: US-10-005-907-1
Perfect score: 3762
Sequence: 1 gagaaacggatcactgtga.....aaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	49	1.3	159	4	US-09-621-976-17182
2	49	1.3	882	2	US-08-909-965C-9
3	49	1.3	1933	4	US-09-920-759-3
4	49	1.3	2851	4	US-09-535-521-1
5	49	1.3	2851	4	US-09-535-521-3
6	48	1.3	191	4	US-09-621-976-9448
7	48	1.3	619	4	US-09-489-847-58
8	48	1.3	1544	4	US-09-187-999-14
9	47	1.2	105	2	US-08-735-381-2
10	47	1.2	105	3	US-09-183-619-1
11	47	1.2	105	3	US-09-201-674-2
12	47	1.2	165	4	US-09-621-976-19195
13	47	1.2	266	4	US-09-621-976-18813
14	47	1.2	326	4	US-09-621-976-16024
15	47	1.2	329	4	US-09-621-976-16012
16	47	1.2	332	4	US-09-621-976-16031
17	47	1.2	332	4	US-09-621-976-16050
18	47	1.2	332	4	US-09-621-976-16053
19	47	1.2	333	4	US-09-621-976-16032
20	47	1.2	333	4	US-09-621-976-16045
21	47	1.2	334	4	US-09-621-976-16044
22	47	1.2	335	4	US-09-621-976-16061
23	47	1.2	336	4	US-09-621-976-16013
24	47	1.2	338	4	US-09-621-976-16041
25	47	1.2	347	4	US-09-621-976-16026
26	47	1.2	357	4	US-09-621-976-16058
27	47	1.2	359	4	US-09-621-976-16008

28	47	1.2	359	4	US-09-621-976-16019
29	47	1.2	362	4	US-09-621-976-16010
30	47	1.2	365	4	US-09-621-976-16042
31	47	1.2	371	4	US-09-621-976-16048
32	47	1.2	615	4	US-09-105-542A-2
33	47	1.2	975	4	US-09-381-488-6
34	47	1.2	1147	1	US-08-665-716-1
35	47	1.2	1454	2	US-08-713-000-7
36	47	1.2	1454	2	US-08-975-316-7
37	47	1.2	1454	3	US-09-211-710-7
38	47	1.2	1454	4	US-09-615-152A-7
39	47	1.2	1454	4	US-09-169-789-7
40	47	1.2	1460	4	US-09-615-192A-95
41	47	1.2	1460	4	US-09-169-789-95
42	47	1.2	1474	2	US-08-975-316-71
43	47	1.2	1474	4	US-09-615-192A-71
44	47	1.2	1474	4	US-09-615-192A-402
45	47	1.2	1474	4	US-09-325-932A-202
46	47	1.2	1474	4	US-09-169-789-71
47	47	1.2	1474	4	US-09-169-789-184
48	47	1.2	2407	3	US-09-370-807-7
49	47	1.2	2407	4	US-09-921-259-7
50	47	1.2	10288	4	US-09-845-917A-28
51	47	1.2	13414	4	US-09-845-917A-27
52	46	1.2	189	4	US-09-621-976-14761
53	46	1.2	193	4	US-09-702-705-102
54	46	1.2	199	4	US-09-736-457-102
55	46	1.2	199	4	US-09-614-124B-102
56	46	1.2	199	4	US-09-671-325-102
57	46	1.2	199	4	US-09-589-184-102
58	46	1.2	213	4	US-09-621-976-16536
59	46	1.2	231	4	US-09-621-976-16317
60	46	1.2	235	4	US-09-621-976-9455
61	46	1.2	242	4	US-09-621-976-16320
62	46	1.2	242	4	US-09-621-976-16324
63	46	1.2	246	4	US-09-621-976-16288
64	46	1.2	249	4	US-09-621-976-16291
65	46	1.2	249	4	US-09-621-976-16292
66	46	1.2	250	4	US-09-621-976-18893
67	46	1.2	259	4	US-09-621-976-16294
68	46	1.2	282	4	US-09-621-976-18648
69	46	1.2	283	4	US-09-621-976-16989
70	46	1.2	356	4	US-09-621-976-16581
71	46	1.2	399	4	US-09-621-976-9484
72	46	1.2	495	4	US-09-220-132-186
73	46	1.2	700	4	US-09-152-060-26
74	46	1.2	776	4	US-09-902-331B-3
75	46	1.2	890	4	US-09-621-976-2725
76	46	1.2	953	2	US-08-197-793-1
77	46	1.2	953	2	US-08-636-176-1
78	46	1.2	953	5	PCT-US95-01618-1
79	46	1.2	997	4	US-09-800-729-14
80	46	1.2	1032	4	US-09-257-178-21
81	46	1.2	1032	4	US-09-464-535-23
82	46	1.2	1071	4	US-09-800-729-45
83	46	1.2	1071	4	US-09-800-729-48
84	46	1.2	1086	4	US-09-800-729-44
85	46	1.2	1138	4	US-09-800-729-47
86	46	1.2	1149	4	US-09-149-476-41
87	46	1.2	1152	4	US-09-439-554-23
88	46	1.2	1192	4	US-09-205-258-25
89	46	1.2	1208	4	US-09-248-335-25
90	46	1.2	1279	3	US-09-164-193-1
91	46	1.2	1315	3	US-09-221-448A-1
92	46	1.2	1315	4	US-09-149-476-208
93	46	1.2	1378	4	US-09-585-173B-7
94	46	1.2	1477	4	US-09-307-794A-220
95	46	1.2	1503	4	US-09-905-125A-220
96	46	1.2	1503	4	US-09-302-775A-220
97	46	1.2	1503	4	US-09-244-111-7
98	46	1.2	1527	4	US-09-820-004-1
99	46	1.2	1606	4	US-09-522-714-11
100	46	1.2	1619	4	

101 46 1.2 1630 4 US-10-162-012-39
102 46 1.2 1699 4 US-09-152-060-19
103 46 1.2 1713 4 US-09-000-062-1
104 46 1.2 1713 4 US-08-945-144A-1
105 46 1.2 1772 2 US-08-960-022-13
106 46 1.2 1790 4 US-09-866-028-31
107 46 1.2 1804 2 US-08-504-459-5
108 46 1.2 1910 3 US-08-674-631-7
109 46 1.2 1958 4 US-08-663-034A-3
110 46 1.2 2013 4 US-09-596-196-3
111 46 1.2 2096 3 US-09-008-481A-10
112 46 1.2 2096 3 US-09-195-666A-16
113 46 1.2 2096 3 US-09-309-592-10
114 46 1.2 2096 4 US-09-635-705-16
115 46 1.2 2096 4 US-09-634-858A-16
116 46 1.2 2096 4 US-08-863-927C-16
117 46 1.2 2205 3 US-08-888-077A-41
118 46 1.2 2223 4 US-09-149-476-24
119 46 1.2 2485 4 US-09-889-463A-9
120 46 1.2 2674 4 US-09-817-180-1
121 46 1.2 2674 4 US-10-003-295-1
122 46 1.2 2744 3 US-09-071-101-1
123 46 1.2 2744 3 US-09-369-618-1
124 46 1.2 2744 3 US-09-369-617-1
125 46 1.2 2822 4 US-09-907-794A-48
126 46 1.2 2822 4 US-09-905-125A-48
127 46 1.2 2822 4 US-09-902-775A-48
128 46 1.2 3138 1 US-07-867-106-4
129 46 1.2 3441 2 US-08-742-783-1
130 46 1.2 3650 4 US-10-072-094-94
131 46 1.2 3722 4 US-10-164-595-9
132 46 1.2 3862 4 US-10-164-595-5
133 46 1.2 3871 2 US-08-599-455B-3
134 46 1.2 3871 3 US-09-069-781B-3
135 46 1.2 3871 4 US-09-137-132-3
136 46 1.2 3871 4 US-08-864-564A-3
137 46 1.2 3871 4 US-09-094-410-3
138 46 1.2 3871 4 US-08-708-123D-3
139 46 1.2 3871 4 US-08-583-153A-3
140 46 1.2 3871 4 US-08-570-142D-3
141 46 1.2 3871 4 US-08-638-524B-3
142 46 1.2 3894 4 US-09-511-625B-3
143 46 1.2 3937 4 US-10-164-595-7
144 46 1.2 3985 4 US-10-164-595-3
145 46 1.2 4456 4 US-09-095-443-1
146 46 1.2 4895 3 US-09-053-866-1
147 46 1.2 4895 4 US-09-479-130-1
148 46 1.2 4895 4 US-09-472-130A-1
149 46 1.2 6671 1 US-08-280-443-1
150 46 1.2 6671 1 US-08-457-459-1

ALIGNMENTS

RESULT 1
US-09-621-976-17182
; Sequence 17182, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET 054PR2
; CURRENT APPLICATION NUMBER: US/09/621, 976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 17182
; LENGTH: 159
; TYPE: DNA
; ORGANISM: Homo sapiens

US-09-621-976-17182

Query Match 1.3%; Score 49; DB 4; Length 159;
Best Local Similarity 100.0%; Pred. No. 3.2e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAA 3762
Db 100 AAATGAAA 148

RESULT 2

US-08-909-965C-9
; Sequence 9, Application US/08909965C
; Patent No. 5936078
; GENERAL INFORMATION:
; APPLICANT: Kuga Tetsuo
; APPLICANT: Nakagawa Satoshi
; APPLICANT: Sakaki Yoshiyuki
; APPLICANT: Zhao Nanding
; APPLICANT: Hashida Hideji
; TITLE OF INVENTION: NOVEL DNA, NOVEL POLYPEPTIDE
; TITLE OF INVENTION: AND NOVEL ANTIBODY
; NUMBER OF SEQUENCES: 17
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: FITZPATRICK, CELLA, HARPER AND SCINTO
; STREET: 277 Park Avenue
; CITY: New York
; STATE: New York
; COUNTRY: U.S.A.
; ZIP: 10172-0194
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/909,965C
; FILING DATE: August 12, 1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: JP 322745/95
; FILING DATE: 12-Dec-1996
; FILING DATE: 12-Dec-1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Lawrence S. Perry
; REGISTRATION NUMBER: 31865
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 212-758-2400
; TELEFAX: 212-758-2982
; TELEX: 236282
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 882 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: human
; IMMEDIATE SOURCE:
; CLONE: F1180
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 97 to 195
; IDENTIFICATION METHOD: by experiment
US-08-909-965C-9

Query Match 1.3%; Score 49; DB 2; Length 882;
Best Local Similarity 100.0%; Pred. No. 2.8e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762
Db 825 AATGAAAAA 873

RESULT 3
US-09-920-759-3
; Sequence 3, Application US/09920759
; Patent No. 6537811
; GENERAL INFORMATION:
; APPLICANT: Brenda F. Baker
; APPLICANT: Susan M. Freier
; TITLE OF INVENTION: ANTISENSE MODULATION OF SAP-1 EXPRESSION
; FILE REFERENCE: RTS-0267
; CURRENT APPLICATION NUMBER: US/09/920,759
; CURRENT FILING DATE: 2001-08-01
; NUMBER OF SEQ ID NOS: 91
; SEQ ID NO 3
; LENGTH: 1933
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (150)....(1445)
US-09-920-759-3

Query Match 1.3%; Score 49; DB 4; Length 1933;
Best Local Similarity 100.0%; Pred. No. 2.6e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762
Db 1878 AATGAAAAA 1926

RESULT 4
US-09-535-521-1
; Sequence 1, Application US/09535521
; Patent No. 6410714
; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 2851
; TYPE: DNA
; ORGANISM: Canis familiaris
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (199)..(1077)
US-09-535-521-1

Query Match 1.3%; Score 49; DB 4; Length 2851;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762
Db 2802 AATGAAAAA 2850

RESULT 5
US-09-535-521-3/c
; Sequence 3, Application US/09535521
; Patent No. 6410714

; GENERAL INFORMATION:
; APPLICANT: Weber, Eric R.
; APPLICANT: McCall, Catherine A.
; TITLE OF INVENTION: NOVEL CANINE LOW AFFINITY IGE RECEPTOR (CANINE CD23)
; FILE REFERENCE: AL-5
; CURRENT APPLICATION NUMBER: US/09/535,521
; CURRENT FILING DATE: 2000-03-24
; EARLIER APPLICATION NUMBER: 60/125,913
; EARLIER FILING DATE: 1999-03-24
; NUMBER OF SEQ ID NOS: 26
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 3
; LENGTH: 2851
; TYPE: DNA
; ORGANISM: Canis familiaris
US-09-535-521-3

Query Match 1.3%; Score 49; DB 4; Length 2851;
Best Local Similarity 100.0%; Pred. No. 2.5e-10;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762
Db 50 AATGAAAAA 2

RESULT 6
US-09-621-976-9448
; Sequence 9448, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9448
; LENGTH: 191
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9448

Query Match 1.3%; Score 48; DB 4; Length 191;
Best Local Similarity 100.0%; Pred. No. 7.9e-10;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762
Db 143 AATGAAAAA 190

RESULT 7
US-09-489-847-58
; Sequence 58, Application US/09489847
; Patent No. 6476195
; GENERAL INFORMATION:
; APPLICANT: Rosen et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: PZ031P1
; CURRENT APPLICATION NUMBER: US/09/489,847
; CURRENT FILING DATE: 2000-01-24
; EARLIER APPLICATION NUMBER: PCT/US99/17130
; EARLIER FILING DATE: 1999-07-29
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; EARLIER APPLICATION NUMBER: 60/095,486
; EARLIER FILING DATE: 1998-08-05
; EARLIER APPLICATION NUMBER: 60/096,319


```

, EARLIER FILING DATE: 1998-08-12
, EARLIER APPLICATION NUMBER: 60/095,454
, EARLIER FILING DATE: 1998-08-06
, EARLIER APPLICATION NUMBER: 60/095,455
, EARLIER FILING DATE: 1998-08-06
, EARLIER APPLICATION NUMBER: 60/095,455
, NUMBER OF SEQ ID NOS: 376
, SOFTWARE: PatentIn Ver. 2.0
, SEQ ID NO 58
, LENGTH: 619
, TYPE: DNA
, ORGANISM: Homo sapiens
, FEATURE:
, NAME/KEY: SITE
, LOCATION: (526)
, OTHER INFORMATION: n equals a,t,g, or c
, FEATURE:
, NAME/KEY: SITE
, LOCATION: (619)
, OTHER INFORMATION: n equals a,t,g, or c
US-09-4889-847-58

```

```

Query Match      1.3%; Score 48; DB 4; Length 619;
Best Local Similarity 100.0%; Pred. No. 7.1e-10;
Matches 48; Conservative 0; Mismatches 0; Indels 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 566 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 613

```

```

RESULT 8
US 09/187-999-14
; Sequence 14, Application US/09187999A
; Patent No. 6482646
; GENERAL INFORMATION:
; APPLICANT: Gindullis, Frank
; APPLICANT: Meier, Iris
; TITLE OF INVENTION: Plant Proteins that Interact with Nuclear Matrix
; TITLE OF INVENTION: Proteins and Function as Transcriptional Activators
; FILE REFERENCE: CL-1321
; CURRENT APPLICATION NUMBER: US/09/187,999A
; CURRENT FILING DATE: 1998-11-06
; NUMBER OF SEQ ID NOS: 39
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 14
; LENGTH: 1544
; TYPE: DNA
; ORGANISM: Lycopersicon esculentum
US-09-187-999-14

```

```

Query Match      1.3%; Score 48; DB 4; Length 1544;
Best Local Similarity 100.0%; Prod. No. 6.5e-10;
Matches 48; Conservative 0; Mismatches 0; Indels 0;

QY 3715 AATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1470 AATCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1517

```

RESULT 9
US-08-735-381-2
Sequence 2, Application US/08735381
Patent No. 5853993
GENERAL INFORMATION:
APPLICANT: Dellinger, Douglas J.
APPLICANT: Dahm, Sueann
APPLICANT: Troll, Mark
TITLE OF INVENTION: SIGNAL ENHANCEMENT METHOD AND KIT
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Hewlett-Packard Company, Legal Dept.,
ADDRESSEE: Intellectual Property
ADDRESSEE: Intellectual Property
STREET: 1501 Page Mill Road, MS 4U-10

CITY: Palo Alto
STATE: California
COUNTRY: USA
ZIP: 94304-1126
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patentin System #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/735,381
FILING DATE: 21-OCT-1996
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Choi, Wendy A.
REGISTRATION NUMBER: 36,697
REFERENCES/DOCKET NUMBER: 10950427-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-857-4125
TELEFAX: 650-852-8063
TELEX: 348-461

```

/ INFORMATION FOR SEQ ID NO: 2:
/
/ SEQUENCE CHARACTERISTICS:
/
/ LENGTH: 105 base pairs
/
/ TYPE: nucleic acid
/
/ STRANDEDNESS: single
/
/ TOPOLOGY: linear
/
/ MOLECULE TYPE: other nucleic acid
/
/ HYPOTHETICAL: NO
/
/ ANTI-SENSE: NO
/
/ US-08-735-381-2

```

```
Query Match      1.2%; Score 47; DB 2; Length 105;
Best Local Similarity 100.0%; Pred. No. 2.1e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 10
US-09-183-619-1
Sequence 1, Application US/09183619
Patent No. 6103474
GENERAL INFORMATION:
APPLICANT: DELLINGER, DOUGLAS J.
APPLICANT: DAHM, SUEANN C.
APPLICANT: ILSLEY, DIANE D.
APPLICANT: ACH, ROBERT A.
APPLICANT: TROLB, MARK A.
TITLE OF INVENTION: HYBRIDIZATION ASSAY SIGNAL ENHANCEMENT
FILE REFERENCE: 10981619-1
CURRENT APPLICATION NUMBER: US/09/183,619
CURRENT FILING DATE: 1998-10-30
EARLIER APPLICATION NUMBER: 08/735,381
EARLIER FILING DATE: 1998-10-21
NUMBER OF SEQ ID NOS: 7
SOFTWARE: PatentIn ver. 2.0
SEQ ID NO 1
LENGTH: 105
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: Description of Artificial Sequence: Tay
US-09-183-619-1

Query Match	1.2%;	Score 47;	DB 3;	Length 105;
Best Local Similarity	100.0%;	Pred. No. 2.1e-09;		
Matches	47;	Conservative	0;	Mismatches 0;
Indels	0			
Qy	3716	ATGAGAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA


```
Db      268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 15
US-09-621-976-16012
; Sequence 16012, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16012
; LENGTH: 329
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16012

Query Match      1.2%; Score 47; DB 4; Length 329;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db      268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 16
US-09-621-976-16031
; Sequence 16031, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16031
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16031

Query Match      1.2%; Score 47; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db      271 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 317

RESULT 17
US-09-621-976-16050
; Sequence 16050, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
```

```
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16050
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16050

Query Match      1.2%; Score 47; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db      268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 18
US-09-621-976-16053
; Sequence 16053, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16053
; LENGTH: 332
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16053

Query Match      1.2%; Score 47; DB 4; Length 332;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db      268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 19
US-09-621-976-16032
; Sequence 16032, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16032
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16032

Query Match      1.2%; Score 47; DB 4; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db      268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314
```

```
RESULT 20
US-09-621-976-16045
; Sequence 16045, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16045
; LENGTH: 333
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16045
Query Match 1.2%; Score 47; DB 4; Length 333;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 21
US-09-621-976-16044
; Sequence 16044, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16044
; LENGTH: 334
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
; LOCATION: 334_
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16044
Query Match 1.2%; Score 47; DB 4; Length 334;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 22
US-09-621-976-16061
; Sequence 16061, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16061
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
US-09-621-976-16061
Query Match 1.2%; Score 47; DB 4; Length 338;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 23
US-09-621-976-16013
; Sequence 16013, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16013
; LENGTH: 336
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16013
Query Match 1.2%; Score 47; DB 4; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 24
US-09-621-976-16041
; Sequence 16041, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16041
; LENGTH: 338
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc_feature
US-09-621-976-16041
Query Match 1.2%; Score 47; DB 4; Length 338;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314
```

```
; LOCATION: 135,197
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16041

Query Match      1.2%; Score 47; DB 4; Length 338;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 270 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 316

RESULT 25
US-09-621-976-16026
; Sequence 16026, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16026
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: 347
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16026

Query Match      1.2%; Score 47; DB 4; Length 347;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 26
US-09-621-976-16058
; Sequence 16058, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16058
; LENGTH: 357
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16058

Query Match      1.2%; Score 47; DB 4; Length 357;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314
```

```
RESULT 27
US-09-621-976-16008
; Sequence 16008, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16008
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16008

Query Match      1.2%; Score 47; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 28
US-09-621-976-16019
; Sequence 16019, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16019
; LENGTH: 359
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16019

Query Match      1.2%; Score 47; DB 4; Length 359;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 29
US-09-621-976-16010
; Sequence 16010, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
```

```
; SOFTWARE: Patent.pm
; SEQ ID NO 16010
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16010

Query Match          1.2%; Score 47; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 30
US-09-621-976-16042
; Sequence 16042, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16042
; LENGTH: 365
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16042

Query Match          1.2%; Score 47; DB 4; Length 365;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 31
US-09-621-976-16048
; Sequence 16048, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 16048
; LENGTH: 371
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16048

Query Match          1.2%; Score 47; DB 4; Length 371;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 269 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 315

; SOFTWARE: Patent.pm
; SEQ ID NO 16010
; LENGTH: 362
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-16010

Query Match          1.2%; Score 47; DB 4; Length 362;
Best Local Similarity 100.0%; Pred. No. 1.9e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 268 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 314

RESULT 32
US-09-105-542A-2
; Sequence 2, Application US/09105542A
; Patent No. 6323329
; GENERAL INFORMATION:
; APPLICANT: Bullerdiel, Jorn
; TITLE OF INVENTION: NUCLEIC ACID SEQUENCES OF GENES OF THE
; FILE REFERENCE: BOHEM4.001C1CP
; CURRENT APPLICATION NUMBER: US/09/105.542A
; CURRENT FILING DATE: 1998-06-26
; PRIOR APPLICATION NUMBER: PCT/DE96/02494
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Homo Sapien
US-09-105-542A-2

Query Match          1.2%; Score 47; DB 4; Length 615;
Best Local Similarity 100.0%; Pred. No. 1.8e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 546 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 592

RESULT 33
US-09-381-488-6
; Sequence 6, Application US/09381488
; Patent No. 6441135
; GENERAL INFORMATION:
; APPLICANT: Kiefer, Michael C.
; APPLICANT: Gibson, Helen L.
; APPLICANT: Fitzpatrick, Paul A.
; APPLICANT: Barr, Philip J.
; TITLE OF INVENTION: A NOVEL Bak BINDING PROTEIN, DNA
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSER: BROWN, THERESA A.
; STREET: 1560 Broadway, Suite 1200
; CITY: Denver
; STATE: CO
; COUNTRY: USA
; ZIP: 80202
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/09/381.488
; FILING DATE: 11-Feb-2000
; CLASSIFICATION: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: BROWN, THERESA A.
; REGISTRATION NUMBER: 32,547
; REFERENCE/DOCKET NUMBER: 4147-15-PUS
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (303) 863-9700
; TELEFAX: (303) 863-0223
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 975 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; FEATURE:
```

```

; NAME/KEY: CDS
; LOCATION: 26...478
; SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-381-488-6
Query Match 1.2%; Score 47; DB 4; Length 975;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 919 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 965
RESULT 34
US-08-665-716-1
; Sequence 1, Application US/08665716
; Patent No. 5789222
; GENERAL INFORMATION:
; APPLICANT: KELLY, ROSEMARIE
; APPLICANT: REGISTER, ELIZABETH A
; APPLICANT: MASUREKAR, PRAKASH S
; TITLE OF INVENTION: P5C REDUCTASE GENE FROM ZALERION
; TITLE OF INVENTION: ARBORICOLA
; NUMBER OF SEQUENCES: 2
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: MERCK & CO., INC.
; STREET: 126 E. LINCOLN AVENUE
; CITY: RAHWAY
; STATE: NEW JERSEY
; COUNTRY: US
; ZIP: 07065-0900
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/665,716
; FILING DATE: 23-JUN-1995
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: KORSEN, ELLIOTT
; REGISTRATION NUMBER: 32,705
; REFERENCE/DOCKET NUMBER: 19453PV
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 908-594-5493
; TELEFAX: 908-594-4720
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1147 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 47..960
US-08-665-716-1
Query Match 1.2%; Score 47; DB 1; Length 1147;
Best Local Similarity 100.0%; Pred. No. 1.7e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1071 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1117
RESULT 35
US-08-713-000-7
; Sequence 7, Application US/08713000
; Patent No. 5850020
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; APPLICANT: Grierson, Alastair W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR THE MODIFICATION OF PLANT LIGNIN CONTENT
; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 15
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Speckman Picard PLLC
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/713,000
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Sleath, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000.1003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-713-000-7
Query Match 1.2%; Score 47; DB 2; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1405 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1451
RESULT 36
US-08-975-316-7
; Sequence 7, Application US/08975316
; Patent No. 5952486
; GENERAL INFORMATION:
; APPLICANT: BLOKSBERG, Leonard N., HAVUKKALA, Ilkka
; APPLICANT: and GRIERSON, Alastair W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR THE MODIFICATION OF PLANT LIGNIN CONTENT
; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/975,316

```

```
; FILING DATE: 1997-11-21
; CLASSIFICATION: 800
; PRIOR APPLICATION NUMBER: 08/713,000
; FILING DATE: September 11, 1996
; FILING DATE: September 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: SLEATH, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000/1003C1
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1454 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
US-08-975-316-7

Query Match 1.2%; Score 47; DB 2; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 37
US-09-211-710-7
; Sequence 7, Application US/09211710A
; Patent No. 6204434
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c3
; CURRENT APPLICATION NUMBER: US/09/211,710A
; CURRENT FILING DATE: 1998-12-14
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-211-710-7

Query Match 1.2%; Score 47; DB 3; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 38
US-09-615-192A-7
; Sequence 7, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1996-09-11
; PRIOR FILING DATE: 1998-10-09
; PRIOR FILING DATE: 1998-10-09
```

```
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-7

Query Match 1.2%; Score 47; DB 4; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 39
US-09-169-789-7
; Sequence 7, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 7
; LENGTH: 1454
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-7

Query Match 1.2%; Score 47; DB 4; Length 1454;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 3762
Db 1405 ATGAAAAA 1451

RESULT 40
US-09-615-192A-95
; Sequence 95, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
```



```
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-95

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e-09; DB 4; Length 1460;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 1.2%; Score 47; DB 4; Length 1460;
| 1411 ATGAAAAA 100.0%; Pred. No. 1.6e-09;
| 1411 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

Db 1411 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 41
US-09-169-789-95
; Sequence 95, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Blockberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 95
; LENGTH: 1460
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-95

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e-09; DB 4; Length 1460;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 1.2%; Score 47; DB 4; Length 1460;
| 1411 ATGAAAAA 100.0%; Pred. No. 1.6e-09;
| 1411 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

Db 1411 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 42
US-08-975-316-71
; Sequence 71, Application US/08975316
; Patent No. 5952486
; GENERAL INFORMATION:
; APPLICANT: BLOCKBERG, Leonard N., HAVUKKALA, Ilkka
; APPLICANT: and GRIERSON, Alastair W.
; TITLE OF INVENTION: MATERIALS AND METHODS FOR
; TITLE OF INVENTION: THE MODIFICATION OF PLANT LIGNIN CONTENT
; NUMBER OF SEQUENCES: 88
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Law Offices of Ann W. Speckman
; STREET: 2601 Elliott Avenue, Suite 4185
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98121
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
```

```
; APPLICATION NUMBER: US/08/975,316
; FILING DATE:
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/713,000
; FILING DATE: September 11, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: SLEATH, Janet
; REGISTRATION NUMBER: 37,007
; REFERENCE/DOCKET NUMBER: 11000/1003C1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 206-269-0565
; TELEFAX: 206-269-0563
; TELEX:
; INFORMATION FOR SEQ ID NO: 71:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1474 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; US-08-975-316-71

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e-09; DB 2; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 1.2%; Score 47; DB 2; Length 1474;
| 1425 ATGAAAAA 100.0%; Pred. No. 1.6e-09;
| 1425 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

Db 1425 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 43
US-09-615-192A-71
; Sequence 71, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Blockberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; TITLE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003C4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 08/975,316
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: US 08/713,000
; PRIOR FILING DATE: 1996-09-11
; PRIOR APPLICATION NUMBER: US 09/169,789
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-71

Query Match
Best Local Similarity 100.0%; Pred. No. 1.6e-09; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA 1.2%; Score 47; DB 4; Length 1474;
| 1425 ATGAAAAA 100.0%; Pred. No. 1.6e-09;
| 1425 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

Db 1425 ATGAAAAA 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 44
US-09-615-192A-402
; Sequence 402, Application US/09615192A
; Patent No. 6410718
; GENERAL INFORMATION:
; APPLICANT: Blockberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
```

```
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c4U
; CURRENT APPLICATION NUMBER: US/09/615,192A
; CURRENT FILING DATE: 2000-07-12
; PRIOR FILING DATE: 1997-11-21
; PRIOR FILING DATE: 1996-09-11
; PRIOR FILING DATE: 1996-09-11
; PRIOR FILING DATE: 1998-10-09
; NUMBER OF SEQ ID NOS: 405
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 402
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-615-192A-402

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA
DB 1425 ATGAAAAA

RESULT 45
US-09-325-932A-202
; Sequence 202, Application US/09325932A
; Patent No. 6451604
; GENERAL INFORMATION:
; APPLICANT: Flinn, Barry
; APPLICANT: Lasham, Annette
; TITLE OF INVENTION: Compositions affecting programmed cell
; FILE OF INVENTION: death and their use in the modification of forestry plant develo
; FILE REFERENCE: 1022
; CURRENT APPLICATION NUMBER: US/09/325,932A
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 206
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 202
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-325-932A-202

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA
DB 1425 ATGAAAAA

RESULT 46
US-09-169-789-71
; Sequence 71, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-169-789-71
```

```
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 71
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-71

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA
DB 1425 ATGAAAAA

RESULT 47
US-09-169-789-184
; Sequence 184, Application US/09169789
; Patent No. 6653528
; GENERAL INFORMATION:
; APPLICANT: Bloksberg, Leonard N.
; APPLICANT: Havukkala, Ilkka
; TITLE OF INVENTION: Materials and Methods for the
; FILE OF INVENTION: Modification of Plant Lignin Content
; FILE REFERENCE: 11000.1003c2
; CURRENT APPLICATION NUMBER: US/09/169,789
; CURRENT FILING DATE: 1998-10-09
; EARLIER APPLICATION NUMBER: US 08/975,316
; EARLIER FILING DATE: 1997-11-21
; EARLIER APPLICATION NUMBER: US 08/713,000
; EARLIER FILING DATE: 1996-09-11
; NUMBER OF SEQ ID NOS: 185
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 184
; LENGTH: 1474
; TYPE: DNA
; ORGANISM: Pinus radiata
US-09-169-789-184

Query Match
Best Local Similarity 100.0%; Score 47; DB 4; Length 1474;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAA
DB 1425 ATGAAAAA

RESULT 48
US-09-370-807-7
; Sequence 7, Application US/09370807
; Patent No. 6297034
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/370,807
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/096,225
; EARLIER FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-370-807-7
```

```
Query Match      1.2%; Score 47; DB 3; Length 2407;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1966 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2012

RESULT 49
US-09-921-259-7
; Sequence 7, Application US/09921259
; Patent No. 6465234
; GENERAL INFORMATION:
; APPLICANT: Cahoon, Rebecca E.
; APPLICANT: Falco, S. Carl
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Sakai, Hajime
; TITLE OF INVENTION: N-End Rule Pathway Enzymes
; FILE REFERENCE: BB-1199
; CURRENT APPLICATION NUMBER: US/09/921,259
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: 60/096,225
; PRIOR FILING DATE: August 12, 1998
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 7
; LENGTH: 2407
; TYPE: DNA
; ORGANISM: Oryza sativa
US-09-921-259-7

Query Match      1.2%; Score 47; DB 4; Length 2407;
Best Local Similarity 100.0%; Pred. No. 1.6e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1966 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2012

RESULT 50
US-09-917A-28/c
; Sequence 28, Application US/09845917A
; Patent No. 6653529
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/845,917A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 28
; LENGTH: 10288
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (8456)
; OTHER INFORMATION: "N is A, C, G, or T"
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificial
; OTHER INFORMATION: Plasmid
US-09-845-917A-28

Query Match      1.2%; Score 47; DB 4; Length 10288;
Best Local Similarity 100.0%; Pred. No. 1.4e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 8083 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 8037

RESULT 51
US-09-845-917A-27/c
; Sequence 27, Application US/09845917A
; Patent No. 6653529
; GENERAL INFORMATION:
; APPLICANT: Bogaert, Thierry
; APPLICANT: Vandekerckhove, Joel
; TITLE OF INVENTION: PROCESSES FOR THE IDENTIFICATION OF COMPOUNDS WHICH
; TITLE OF INVENTION: CONTROL CELL BEHAVIOUR, THE COMPOUNDS IDENTIFIED AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS CONTAINING THEM AND THEIR
; TITLE OF INVENTION: USE IN THE CONTROL OF CELL BEHAVIOUR
; FILE REFERENCE: P/14-1
; CURRENT APPLICATION NUMBER: US/09/845,917A
; CURRENT FILING DATE: 2001-04-30
; NUMBER OF SEQ ID NOS: 58
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 27
; LENGTH: 13414
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (11582)
; OTHER INFORMATION: "N is A, G, C or T"
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: artificial
; OTHER INFORMATION: Plasmid
US-09-845-917A-27

Query Match      1.2%; Score 47; DB 4; Length 13414;
Best Local Similarity 100.0%; Pred. No. 1.3e-09;
Matches 47; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3716 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 11209 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 11163

RESULT 52
US-09-621-976-14761
; Sequence 14761, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 14761
; LENGTH: 189
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-14761

Query Match      1.2%; Score 46; DB 4; Length 189;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 114 TGAATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 159
```

[illegible][illegible]

APPLICANT: Wang, Tongtong
APPLICANT: Baugur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
APPLICANT: Fan, Liqun
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
FILE REFERENCE: 210121.478C12
CURRENT APPLICATION NUMBER: US/09/671.325
CURRENT FILING DATE: 2000-09-26
NUMBER OF SEQ ID NOS: 1825
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 102
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapien
US-09-671-325-102

Query Match 1.2%; Score 46; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAAGAAAAA
Db 154 TGAAGAAAAA

RESULT 58
US-09-589-184-102
Sequence 102, Application US/09589184
Patent No. 6686447
GENERAL INFORMATION:
APPLICANT: Wang, Tongtong
APPLICANT: Baugur, Chaitanya S.
APPLICANT: Lodes, Michael A.
APPLICANT: Fanger, Gary
APPLICANT: Vedvick, Tom
APPLICANT: Carter, Darrick
APPLICANT: Retter, Marc
APPLICANT: Mannion, Jane
TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
FILE REFERENCE: 210121.478C8
CURRENT APPLICATION NUMBER: US/09/589,184
CURRENT FILING DATE: 2000-06-05
NUMBER OF SEQ ID NOS: 827
SOFTWARE: FastSEQ for Windows Version 3.0
SEQ ID NO 102
LENGTH: 199
TYPE: DNA
ORGANISM: Homo sapien
US-09-589-184-102

Query Match 1.2%; Score 46; DB 4; Length 199;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAAGAAAAA
Db 154 TGAAGAAAAA

RESULT 59
US-09-621-976-16536
Sequence 16536, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.

APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16536
LENGTH: 213
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-16536

Query Match 1.2%; Score 46; DB 4; Length 213;
Best Local Similarity 100.0%; Pred. No. 4.9e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAAGAAAAA
Db 156 TGAAGAAAAA

RESULT 60
US-09-621-976-16317
Sequence 16317, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 16317
LENGTH: 231
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: 231
OTHER INFORMATION: n=a, g, c or t
US-09-621-976-16317

Query Match 1.2%; Score 46; DB 4; Length 231;
Best Local Similarity 100.0%; Pred. No. 4.8e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TGAAGAAAAA
Db 160 TGAAGAAAAA

RESULT 61
US-09-621-976-9455
Sequence 9455, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 9455
LENGTH: 235
TYPE: DNA
ORGANISM: Homo sapiens


```

; LOCATION: 230..231
; OTHER INFORMATION: n=a, g, c o r t
US-09-621-976-16561

Query Match      1.2%; Score 46; DB 4; Length 356;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TGAAGAAAAA 3762
Db 308 TGAAGAAAAA 353

RESULT 72
US-09-621-976-9484
; Sequence 9484, Application US/09621976
; Patent NO. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 9484
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-9484

Query Match      1.2%; Score 46; DB 4; Length 399;
Best Local Similarity 100.0%; Pred. No. 4.6e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TGAAGAAAAA 3762
Db 349 TGAAGAAAAA 394

RESULT 73
US-09-220-132-186/c
; Sequence 186, Application US/09220132
; Patent NO. 6506607
; GENERAL INFORMATION:
; APPLICANT: Shyjan, Andrew W.
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR THE IDENTIFICATION AND ASSESSMENT
; FILE REFERENCE: 07334-074001
; CURRENT APPLICATION NUMBER: US/09/220.132
; CURRENT FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: US 60/079,303
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: US 60/068,821
; PRIOR FILING DATE: 1997-12-24
; NUMBER OF SEQ ID NOS: 191
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 186
; LENGTH: 495
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-220-132-186

Query Match      1.2%; Score 46; DB 4; Length 495;
Best Local Similarity 100.0%; Pred. No. 4.5e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TGAAGAAAAA 3762
Db 64 TGAAGAAAAA 19

```



```

1  COUNTRY: USA
2  ZIP: 02109
3  COMPUTER READABLE FORM:
4  MEDIUM TYPE: Floppy disk
5  COMPUTER: IBM PC compatible
6  OPERATING SYSTEM: PC-DOS/MS-DOS
7  SOFTWARE: ASCII text
8  CURRENT APPLICATION DATA:
9  APPLICATION NUMBER: US/08/197,793
10 FILING DATE:
11 CLASSIFICATION: 530
12 PRIOR APPLICATION DATA:
13 APPLICATION NUMBER: 07/688,019;
14 FILING DATE: 19-APR-1991
15 APPLICATION NUMBER: 08/004,199
16 FILING DATE: 13-JAN-1993
17 ATTORNEY/AGENT INFORMATION:
18 NAME: DeConti, Giulio A., Jr.
19 REGISTRATION NUMBER: 31,503
20 REFERENCE/DOCKET NUMBER: BBI-006CNCNP
21 TELECOMMUNICATION INFORMATION:
22 TELEPHONE: (617) 227-7400
23 TELEFAX: (617) 227-5941
24 INFORMATION FOR SEQ ID NO: 1:
25 SEQUENCE CHARACTERISTICS:
26 LENGTH: 953 base pairs
27 TYPE: nucleic acid
28 STRANDEDNESS: single
29 TOPOLOGY: linear
30 MOLECULE TYPE: cDNA
31 FEATURE:
32 NAME/KEY: CDS
33 LOCATION: 64..681
34 US-08-197-793-1
35
36 Query Match 1.2%; Score 46; DB 1; Length 953;
37 Best Local Similarity 100.0%; Pred.No. 4.2e-09;
38 Matches 46; Conservative 0; Mismatches 0; Indels 0
39
40 QY 3717 TCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
41 DB 905 TCGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 950
42
43 RESULT 78
44 US-08-636-176-1
45 Sequence 1, Application US/08636176
46 Patent No. 5846822
47 GENERAL INFORMATION:
48 APPLICANT: Meuer, S.
49 APPLICANT: Schraven, B.
50 APPLICANT: Schoenaut, D.
51 APPLICANT: Ratnosky, S.
52 TITLE OF INVENTION: pp32: A Newly Identified CD45-Associated
53 TITLE OF INVENTION: Protein
54 NUMBER OF SEQUENCES: 20
55 CORRESPONDENCE ADDRESS:
56 ADDRESSEE: LAHIVE & COCKFIELD
57 STREET: 60 STATE STREET, SUITE 510
58 CITY: BOSTON
59 STATE: MASSACHUSETTS
60 COUNTRY: USA
61 ZIP: 02109
62 COMPUTER READABLE FORM:
63 MEDIUM TYPE: Floppy disk
64 COMPUTER: IBM PC compatible
65 OPERATING SYSTEM: PC-DOS/MS-DOS
66 SOFTWARE: ASCII text
67 CURRENT APPLICATION DATA:
68 APPLICATION NUMBER: US/08/636,176
69 FILING DATE:
70 CLASSIFICATION: 530
71 PRIOR APPLICATION DATA:

```

```
; APPLICATION NUMBER: 07/688,019; 08/004,199
; FILING DATE: 19-APR-1991; 13-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-006CNCPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 953 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64...681
; US-08-636-176-1

Query Match 1.2%; Score 46; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 905 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 950

RESULT 79
PCT-US95-01618-1
; Sequence 1, Application PC/TUS9501618
; GENERAL INFORMATION:
; APPLICANT:
; TITLE OF INVENTION: PB32: A Newly Identified CD45-Associated
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 60 STATE STREET, SUITE 510
; CITY: BOSTON
; STATE: MASSACHUSETTS
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: ASCII text
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01618
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: USSN 08/197,793
; ATTORNEY/AGENT INFORMATION:
; NAME: DeConti, Giulio A., Jr.
; REGISTRATION NUMBER: 31,503
; REFERENCE/DOCKET NUMBER: BBI-006CPCPP
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 227-5941
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 953 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 64...681

APPLICATION NUMBER: 07/688,019; 08/004,199
FILING DATE: 19-APR-1991; 13-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: DeConti, Giulio A., Jr.
REGISTRATION NUMBER: 31,503
REFERENCE/DOCKET NUMBER: BBI-006CNCPP
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 227-7400
TELEFAX: (617) 227-5941
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 953 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 64...681
US-08-636-176-1

Query Match 1.2%; Score 46; DB 2; Length 953;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 905 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 950

RESULT 80
US-09-800-729-14
; Sequence 14, Application US/09800729
; Patent No. 6605592
; GENERAL INFORMATION:
; APPLICANT: Ni et al.
; TITLE OF INVENTION: 32 Human secreted proteins
; FILE REFERENCE: PZ044P1
; CURRENT APPLICATION NUMBER: US/09/800,729
; CURRENT FILING DATE: 2001-03-08
; PRIOR APPLICATION NUMBER: PCT/US00/26013
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 60/155,709
; PRIOR FILING DATE: 1999-09-24
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 14
; LENGTH: 997
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-800-729-14

Query Match 1.2%; Score 46; DB 4; Length 997;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 951 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 996

RESULT 81
US-09-257-179-21
; Sequence 21, Application US/09257179
; Patent No. 6410709
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: 29 Human Secreted Proteins
; FILE REFERENCE: PZ015P1
; CURRENT APPLICATION NUMBER: US/09/257,179
; CURRENT FILING DATE: 1999-02-25
; EARLIER APPLICATION NUMBER: PCT/US98/17709
; EARLIER FILING DATE: 1998-08-27
; EARLIER APPLICATION NUMBER: 60/056,270
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,271
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,247
; EARLIER FILING DATE: 1997-08-29
; EARLIER APPLICATION NUMBER: 60/056,073
; EARLIER FILING DATE: 1997-08-29
; NUMBER OF SEQ ID NOS: 128
; SOFTWARE: Patent In Ver. 2.0
; SEQ ID NO 21
; LENGTH: 1032
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-257-179-21

Query Match 1.2%; Score 46; DB 4; Length 1032;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

.....

THE UNIVERSITY OF CHICAGO

```
; NUMBER OF SEQ ID NOS: 217
; SOFTWARE: Patentin Ver: 2.0
; SEQ ID NO 47
; LENGTH: 1149
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-800-729-47

Query Match
Best Local Similarity 1.2%; Score 46; DB 4; Length 1149;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3717 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1063 TCAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1108

RESULT 87
US-09-149-476-41
; Sequence 41, Application US/09149476
; Patent No. 6420526
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: 186 Human Secreted proteins
; FILE REFERENCE: P2002P1
; CURRENT APPLICATION NUMBER: US/09/149,476
; CURRENT FILING DATE: 1998-09-08
; EARLIER APPLICATION NUMBER: PCT/US98/04493
; EARLIER FILING DATE: 1998-03-06
; EARLIER APPLICATION NUMBER: 60/040,162
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,333
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/038,621
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,626
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,334
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,336
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/040,163
; EARLIER FILING DATE: 1997-03-07
; EARLIER APPLICATION NUMBER: 60/047,600
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,615
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,597
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,502
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,633
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,583
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,617
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,618
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,503
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,592
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,581
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,584
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,500
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,587
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,492
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,598
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,613
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,582
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,596
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,612
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,632
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/047,601
; EARLIER FILING DATE: 1997-05-23
; EARLIER APPLICATION NUMBER: 60/043,580
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,568
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,314
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,569
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,311
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,671
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,674
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,669
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,312
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,313
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,672
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/043,315
; EARLIER FILING DATE: 1997-04-11
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/056,886
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,877
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,889
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,893
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,630
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,878
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,662
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,872
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,882
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,637
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,903
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,888
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,879
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,880
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,894
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,911
; EARLIER FILING DATE: 1997-08-22
; EARLIER APPLICATION NUMBER: 60/056,636
```

```

Query Match          1.2%; Score 46; DB 4; Length 1153;
Best Local Similarity 100.0%; Pred. No. 4.2e-09;
Matches 46; Conservative 0; Mismatches 0; Indels

QY 3717 TGAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1094 TGAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1139

RESULT 88
US-09-439-554-23
; Sequence 23, Application US/09439554
; Patent No. 6479733
; GENERAL INFORMATION:
; APPLICANT: Rafalski, Jan Antoni
; APPLICANT: Odell, Joan T.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; APPLICANT: Kinney, Anthony J.
; APPLICANT: Famodu, Omolayo O.
; TITLE OF INVENTION: STEROL METABOLISM ENZYMES
; FILE REFERENCE: BB1114 US NA
; CURRENT APPLICATION NUMBER: US/09/439,554
; CURRENT FILING DATE: 1999-11-12
; EARLIER APPLICATION NUMBER: 60/108,351
; EARLIER FILING DATE: 1998-No. 6479733ember-13
; NUMBER OF SEQ ID NOS: 30
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 23
; LENGTH: 1192
; TYPE: DNA
; ORGANISM: Glycine max
US-09-439-554-23

Query Match          1.2%; Score 46; DB 4; Length 1192;
Best Local Similarity 100.0%; Pred. No. 4.1e-09;
Matches 46; Conservative 0; Mismatches 0; Indels

QY 3717 TGAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1103 TGAIAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1148

RESULT 89
US-09-205-258-25
; Sequence 25, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; FILE REFERENCE: P2007p1
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884

```


; TITLE OF INVENTION: NOVEL CSAPT PROTEIN MOLECULES AND USES THEREFOR (AS AMENDED)

; FILE REFERENCE: MNI-051DV1

; CURRENT APPLICATION NUMBER: US/09/221,448A

; CURRENT FILING DATE: 1998-12-28

; NUMBER OF SEQ ID NOS: 23

; SOFTWARE: Patent in Ver. 2.0

; SEQ ID NO 1

; LENGTH: 1315

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: CDS

; LOCATION: (249)..(767)

; OTHER INFORMATION: 'n' at position 1315 may be any nucleotide

US-09-221-448A-1

Query Match 1.2%; Score 46; DB 4; Length 1315;
Best Local Similarity 100.0%; Fred. No. 4.1e-03;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3717 TCAA 3762

Db 1245 TCAA 1290

RESULT 93

US-09-149-476-208

; Sequence 208, Application US/09149476

; Patent No. 6420526

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: 186 Human Secreted proteins

; FILE REFERENCE: P2002P1

; CURRENT APPLICATION NUMBER: US/09/149,476

; CURRENT FILING DATE: 1998-09-08

; EARLIER APPLICATION NUMBER: PCT/US98/04493

; EARLIER FILING DATE: 1998-03-06

; EARLIER APPLICATION NUMBER: 60/040,162

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,333

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/038,621

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,626

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,334

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,336

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/040,163

; EARLIER FILING DATE: 1997-03-07

; EARLIER APPLICATION NUMBER: 60/047,600

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,615

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,597

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,502

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,633

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,583

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,617

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,618

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,503

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,592

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,581

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,584

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,500

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,587

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,492

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,598

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,584

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,500

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,587

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,492

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,598

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,613

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,582

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,596

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,612

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,632

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/047,601

; EARLIER FILING DATE: 1997-05-23

; EARLIER APPLICATION NUMBER: 60/043,580

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,568

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,314

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,569

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,311

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,671

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,674

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,669

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,312

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,313

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,672

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/043,315

; EARLIER FILING DATE: 1997-04-11

; EARLIER APPLICATION NUMBER: 60/048,974

; EARLIER FILING DATE: 1997-08-06

; EARLIER APPLICATION NUMBER: 60/056,886

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,877

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,889

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,893

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,630

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,878

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,662

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,872

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,882

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,637

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,903

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,888

; EARLIER FILING DATE: 1997-08-22

; EARLIER APPLICATION NUMBER: 60/056,879


```

; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-902-775A-220

Query Match      1.2%; Score 46; DB 4; Length 1503;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3717 TGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3762
Db      1445 TGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1490

RESULT 98
US-09-244-111-7
; Sequence 7, Application US/09244111
; Patent No. 656498
; GENERAL INFORMATION:
; APPLICANT: NI, et al.
; TITLE OF INVENTION: Human Serine Protease and Serpin Polypeptides
; FILE REFERENCE: PF391
; CURRENT APPLICATION NUMBER: US/09/244,111
; CURRENT FILING DATE: 1999-02-04
; EARLIER APPLICATION NUMBER: 60/073,961
; EARLIER FILING DATE: 1998-02-06
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 7
; LENGTH: 1527
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (67)..(1335)
US-09-244-111-7

Query Match      1.2%; Score 46; DB 4; Length 1527;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3717 TGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3762
Db      1474 TGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1519

RESULT 99
US-09-820-004-1
; Sequence 1, Application US/09820004
; Patent No. 6649385
; GENERAL INFORMATION:
; APPLICANT: WEI, Ming-Hui et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CLO01201
; CURRENT APPLICATION NUMBER: US/09/820,004
; CURRENT FILING DATE: 2001-03-29
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 1606
; TYPE: DNA
; ORGANISM: Human
US-09-820-004-1

Query Match      1.2%; Score 46; DB 4; Length 1606;
Best Local Similarity 100.0%; Pred. No. 4e-09;
Matches 46; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3717 TGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 3762
Db      1532 TGAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 1577

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 08:22:52 ; Search time 1469 Seconds

(without alignments)
11546.259 Million cell updates/sec

Title: US-10-005-907-1

Perfect score: 3762

Sequence: 1 gagaacgggtcactgtga.....aaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC

Gapex 60.0

Searched: 2907579 seqs, 225431464 residues

Word size : 0

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 150 summaries

Database : Published Applications NA.*

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq.*
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq.*
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq.*
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3762	100.0	3762	15	US-10-005-907-1
2	50	1.3	261	13	US-10-424-599-141898
3	50	1.3	1702	9	US-09-925-300-395
4	50	1.3	3602	15	US-10-171-311-67
5	49	1.3	476	14	US-10-001-843-87
6	49	1.3	606	13	US-10-424-599-35023
7	49	1.3	1057	15	US-10-106-998-741
8	49	1.3	1372	9	US-09-925-301-10
9	49	1.3	1393	15	US-10-198-846-9741
10	49	1.3	1417	13	US-10-424-599-11755
11	49	1.3	1423	13	US-10-424-599-106798
12	49	1.3	1429	15	US-10-007-926A-34
13	49	1.3	1998	9	US-09-764-903-19
14	49	1.3	1998	15	US-10-091-391-19

49	1.3	2143	16	US-10-264-049-740	Sequence 740, Appl
49	1.3	2713	15	US-10-436-523-1	Sequence 1, Appl
49	1.3	2713	15	US-10-436-523-31	Sequence 31, Appl
49	1.3	2713	15	US-10-436-523-32	Sequence 32, Appl
49	1.3	2713	15	US-10-436-523-33	Sequence 33, Appl
49	1.3	2713	15	US-10-436-523-34	Sequence 34, Appl
49	1.3	2944	15	US-10-128-714-1558	Sequence 1558, Ap
49	1.3	3362	15	US-10-128-714-6558	Sequence 6558, Ap
49	1.3	4237	9	US-09-745-763-20	Sequence 20, Appl
49	1.3	4944	15	US-10-128-714-558	Sequence 558, App
49	1.3	5362	15	US-10-128-714-5558	Sequence 5558, Ap
48	1.3	337	9	US-09-960-352-6036	Sequence 6036, Ap
48	1.3	348	13	US-10-085-783A-16208	Sequence 16208, A
48	1.3	348	16	US-10-242-535A-16208	Sequence 16208, A
48	1.3	372	13	US-10-424-599-47139	Sequence 47139, A
48	1.3	401	10	US-09-918-985-16319	Sequence 16319, A
48	1.3	462	13	US-10-424-599-54555	Sequence 54555, A
48	1.3	544	13	US-10-424-599-46783	Sequence 46783, A
48	1.3	619	13	US-10-351-334-58	Sequence 58, Appl
48	1.3	751	13	US-10-424-599-95989	Sequence 95989, A
48	1.3	865	9	US-09-770-445-579	Sequence 579, App
48	1.3	1837	13	US-10-219-535-39	Sequence 39, Appl
48	1.3	1837	13	US-10-232-230-39	Sequence 39, Appl
48	1.3	1837	13	US-10-232-230-39	Sequence 39, Appl
48	1.3	1837	15	US-10-227-884-39	Sequence 39, Appl
48	1.3	1837	15	US-10-230-163-39	Sequence 39, Appl
48	1.3	1837	15	US-10-230-338-39	Sequence 39, Appl
48	1.3	1837	15	US-10-218-631-39	Sequence 39, Appl
48	1.3	1837	15	US-10-230-414-39	Sequence 39, Appl
48	1.3	1837	15	US-10-216-159A-39	Sequence 39, Appl
48	1.3	1837	15	US-10-218-849-39	Sequence 39, Appl
48	1.3	1837	15	US-10-227-873-39	Sequence 39, Appl
48	1.3	1837	15	US-10-227-883-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-076-39	Sequence 39, Appl
48	1.3	1837	15	US-10-230-434-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-003-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-075-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-464-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-466-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-479-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-481-39	Sequence 39, Appl
48	1.3	1837	15	US-10-230-260-39	Sequence 39, Appl
48	1.3	1837	15	US-10-232-231-39	Sequence 39, Appl
48	1.3	1837	15	US-10-232-233-39	Sequence 39, Appl
48	1.3	1837	15	US-10-216-185-39	Sequence 39, Appl
48	1.3	1837	15	US-10-218-956-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-468-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-478-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-536-39	Sequence 39, Appl
48	1.3	1837	15	US-10-233-205-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-072-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-470-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-474-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-524-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-528-39	Sequence 39, Appl
48	1.3	1837	15	US-10-227-880-39	Sequence 39, Appl
48	1.3	1837	15	US-10-227-881-39	Sequence 39, Appl
48	1.3	1837	15	US-10-227-882-39	Sequence 39, Appl
48	1.3	1837	15	US-10-230-436-39	Sequence 39, Appl
48	1.3	1837	15	US-10-232-223-39	Sequence 39, Appl
48	1.3	1837	15	US-10-232-225-39	Sequence 39, Appl
48	1.3	1837	15	US-10-232-227-39	Sequence 39, Appl
48	1.3	1837	15	US-10-232-229-39	Sequence 39, Appl
48	1.3	1837	15	US-10-232-234-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-060-39	Sequence 39, Appl
48	1.3	1837	15	US-10-216-160-39	Sequence 39, Appl
48	1.3	1837	15	US-10-216-162-39	Sequence 39, Appl
48	1.3	1837	15	US-10-216-164-39	Sequence 39, Appl
48	1.3	1837	15	US-10-216-167-39	Sequence 39, Appl
48	1.3	1837	15	US-10-216-168-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-065-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-071-39	Sequence 39, Appl
48	1.3	1837	15	US-10-219-074-39	Sequence 39, Appl

88 1.3 1837 15 US-10-219-077-39
89 1.3 1837 15 US-10-219-465-39
90 1.3 1837 15 US-10-219-467-39
91 1.3 1837 15 US-10-219-469-39
92 1.3 1837 15 US-10-219-471-39
93 1.3 1837 15 US-10-219-473-39
94 1.3 1837 15 US-10-219-476-39
95 1.3 1837 15 US-10-219-482-39
96 1.3 1837 15 US-10-227-874-39
97 1.3 1837 15 US-10-227-876-39
98 1.3 1837 15 US-10-229-878-39
99 1.3 1837 15 US-10-229-974-39
100 1.3 1837 15 US-10-230-024-39
101 1.3 1837 15 US-10-230-113-39
102 1.3 1837 15 US-10-230-183-39
103 1.3 1837 15 US-10-230-234-39
104 1.3 1837 15 US-10-230-306-39
105 1.3 1837 15 US-10-230-426-39
106 1.3 1837 15 US-10-230-427-39
107 1.3 1837 15 US-10-230-433-39
108 1.3 1837 15 US-10-230-435-39
109 1.3 1837 15 US-10-230-438-39
110 1.3 1837 15 US-10-232-222-39
111 1.3 1837 15 US-10-219-070-39
112 1.3 1837 15 US-10-219-472-39
113 1.3 1837 15 US-10-219-527-39
114 1.3 1837 15 US-10-227-877-39
115 1.3 1837 15 US-10-216-166-39
116 1.3 1837 15 US-10-218-612-39
117 1.3 1837 15 US-10-216-163-39
118 1.3 1837 15 US-10-218-765-39
119 1.3 1837 15 US-10-219-063-39
120 1.3 1837 15 US-10-219-066-39
121 1.3 1837 15 US-10-219-067-39
122 1.3 1837 15 US-10-219-068-39
123 1.3 1837 15 US-10-219-069-39
124 1.3 1837 15 US-10-219-073-39
125 1.3 1837 15 US-10-219-475-39
126 1.3 1837 15 US-10-219-480-39
127 1.3 1837 15 US-10-219-483-39
128 1.3 1837 15 US-10-219-525-39
129 1.3 1837 15 US-10-219-526-39
130 1.3 1837 15 US-10-219-530-39
131 1.3 1837 15 US-10-219-531-39
132 1.3 1837 15 US-10-219-532-39
133 1.3 1837 15 US-10-219-533-39
134 1.3 1837 15 US-10-230-437-39
135 1.3 1837 15 US-10-232-228-39
136 1.3 1837 15 US-10-232-226-39
137 1.3 1837 15 US-10-230-130-39
138 1.3 1837 15 US-10-424-599-12512
139 1.3 1837 15 US-10-219-535-7
140 1.3 1837 15 US-10-232-230-7
141 1.3 1837 15 US-10-232-224-7
142 1.3 1837 15 US-10-227-884-7
143 1.3 1837 15 US-10-230-163-7
144 1.3 1837 15 US-10-230-338-7
145 1.3 1837 15 US-10-230-331-7
146 1.3 1837 15 US-10-218-631-7
147 1.3 1837 15 US-10-230-414-7
148 1.3 1837 15 US-10-216-159A-7
149 1.3 1837 15 US-10-218-849-7
150 1.3 1837 15 US-10-227-873-7
151 1.3 1837 15 US-10-227-883-7

ALIGNMENTS

RESULT 1
US-10-005-907-1
; Sequence 1, Application US/10005907
; Publication NO. US2003016681A1
; GENERAL INFORMATION:

; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US2003016681A1, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST CELL
; TITLE OF INVENTION: ACTIVATION
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005,907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 3762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(432)
; OTHER INFORMATION:
US-10-005-907-1
Query Match 100.0%; Score 3762; DB 15; Length 3762;
Best Local Similarity 100.0%; Pred No. 0;
Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GAGAAACCGAGTCACCTGTGAAAGATGGAAATTTATCTCTCGGAAATCTAGTTGCCTG 60
Db 1 GAGAAACCGAGTCACCTGTGAAAGATGGAAATTTATCTCTCGGAAATCTAGTTGCCTG 60
Qy 61 GGAGAGATCAAGAGAGCCAG 120
Db 61 GGAGAGATCAAGAGAGCCAG 120
Qy 121 ATGACTACATTTTGAAGAGAAATCTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 180
Db 121 ATGACTACATTTTGAAGAGAAATCTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 180
Qy 181 ACTTCTAATCAG 240
Db 181 ACTTCTAATCAG 240
Qy 241 AATCAGATCCCCATCAGAGATCCTCCCTGAGCTCCATGATGATGATGATGATGATGATGAT 300
Db 241 AATCAGATCCCCATCAGAGATCCTCCCTGAGCTCCATGATGATGATGATGATGATGATGAT 300
Qy 301 GACTCCCTCAAG 360
Db 301 GACTCCCTCAAG 360
Qy 361 CTTAGGAGCTTCTGTAGTAGGCTTGTCTGACCCATGAGCATGATGATGATGATGATGATGAT 420
Db 361 CTTAGGAGCTTCTGTAGTAGGCTTGTCTGACCCATGAGCATGATGATGATGATGATGATGAT 420
Qy 421 TTTCACACATAAATCTCAAGCTGCTTTATACCTTCAGCAATGAGCAATGAGCAATGAGCAAT 480
Db 421 TTTCACACATAAATCTCAAGCTGCTTTATACCTTCAGCAATGAGCAATGAGCAATGAGCAAT 480
Qy 481 TAGCAGACTCTGCGAGAGTTGTTTCACTGAGCAGTGCATGAAACATTCCTTTCTGGCTA 540
Db 481 TAGCAGACTCTGCGAGAGTTGTTTCACTGAGCAGTGCATGAAACATTCCTTTCTGGCTA 540
Qy 541 AAGTTTGAAGATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 600
Db 541 AAGTTTGAAGATTTATTTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 600
Qy 601 TGGCTTAGGTAATCATAGAAATTCAGCAATGAGCAATGAGCAATGAGCAATGAGCAATGAGCAAT 660
Db 601 TGGCTTAGGTAATCATAGAAATTCAGCAATGAGCAATGAGCAATGAGCAATGAGCAATGAGCAAT 660
Qy 661 TTGTAAAGTTTGAAG 720
Db 661 TTGTAAAGTTTGAAG 720

QY 721 TGAATAAATTTCTAATCCCTGACTAACTGAATGGACCCCTCTTAGCCAAAGAGA 780
Db 721 TGAATAAATTTCTAATCCCTGACTAACTGAATGGACCCCTCTTAGCCAAAGAGA 780
QY 781 CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGAAAGGAGGTGAGACACACC 840
Db 781 CCTCAGATGAACCTGAAAGACTGAATTTCTGGCCATGATAGAAAGGAGGTGAGACACACC 840
QY 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAACAGTGACCGAGGTAGTCAATAGACT 900
Db 841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAACAGTGACCGAGGTAGTCAATAGACT 900
QY 901 GATGAATAGACTGATTTGGCAATAAGAGTCCCAATTCCAACTGACTCTGGGTAGAT 960
Db 901 GATGAATAGACTGATTTGGCAATAAGAGTCCCAATTCCAACTGACTCTGGGTAGAT 960
QY 961 CACACACTGCTGAGGATTCATCTATGAGACTTTGTCTACATACACAGAGACTTGGTT 1020
Db 961 CACACACTGCTGAGGATTCATCTATGAGACTTTGTCTACATACACAGAGACTTGGTT 1020
QY 1021 TCCACAACCCCTTATTTTAACTAAAGCAATTTCTTCTACTGACTCTTAAAGTCTTTAGA 1080
Db 1021 TCCACAACCCCTTATTTTAACTAAAGCAATTTCTTCTACTGACTCTTAAAGTCTTTAGA 1080
QY 1081 CAAAGCTTAACCTTTCAACCAATGGCAATGCAACAACTTTGAATCTACCTATGACCT 1140
Db 1081 CAAAGCTTAACCTTTCAACCAATGGCAATGCAACAACTTTGAATCTACCTATGACCT 1140
QY 1141 GTAAGCTCTCTCTGCTTCAAGACTTCTGCTCTTTTAACTGAAACCGATGCACTTTCCA 1200
Db 1141 GTAAGCTCTCTCTGCTTCAAGACTTCTGCTCTTTTAACTGAAACCGATGCACTTTCCA 1200
QY 1201 TTTAATGATTTAGTCTTTGCTTGTAACTCTCTCTCTCCCTTAAATGTATAAAAGTAAACG 1260
Db 1201 TTTAATGATTTAGTCTTTGCTTGTAACTCTCTCTCTCCCTTAAATGTATAAAAGTAAACG 1260
QY 1261 GTGACCTGACCACTCAGGACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
Db 1261 GTGACCTGACCACTCAGGACACTTTCTCAGGACCTCTCTGAGAGTGTATCCAGGCCAT 1320
QY 1321 GGTAAAGTCATGTTGGCTCAGAACTCAACCTCTTTTAAATATTTTACAGAAATTTGGTTTGG 1380
Db 1321 GGTAAAGTCATGTTGGCTCAGAACTCAACCTCTTTTAAATATTTTACAGAAATTTGGTTTGG 1380
QY 1381 TTACCAATAGTCTCCCAAAATATATGTCCAAGATCTTCAATTCAGGCTCTCACCACCA 1440
Db 1381 TTACCAATAGTCTCCCAAAATATATGTCCAAGATCTTCAATTCAGGCTCTCACCACCA 1440
QY 1441 AATTTCAAATGCCAATCTCCCATCAATTTACTATTTTCACTTTGAGGTGTAACTA 1500
Db 1441 AATTTCAAATGCCAATCTCCCATCAATTTACTATTTTCACTTTGAGGTGTAACTA 1500
QY 1501 CTCGAATAAAGTGTAAAGCAGTGACAGACCTTTGCTTAACTGACATTTACTTCAAT 1560
Db 1501 CTCGAATAAAGTGTAAAGCAGTGACAGACCTTTGCTTAACTGACATTTACTTCAAT 1560
QY 1561 TTTTCTTTTCTAGTACTGATATTTTGCATATAAATTTGAGTAAATAGTTTCAAAAT 1620
Db 1561 TTTTCTTTTCTAGTACTGATATTTTGCATATAAATTTGAGTAAATAGTTTCAAAAT 1620
QY 1621 TAATAGTTTGTGACATTTGGCTTTCTGAGAAGAGAAATGAAAGTGTCAAAATAAATA 1680
Db 1621 TAATAGTTTGTGACATTTGGCTTTCTGAGAAGAGAAATGAAAGTGTCAAAATAAATA 1680
QY 1681 AAGATGAATGAACATATATATGTCATTTTCTCAATTTTCTAGTCAACAGAGATC 1740
Db 1681 AAGATGAATGAACATATATATGTCATTTTCTCAATTTTCTAGTCAACAGAGATC 1740
QY 1741 GAAGGATTTCTGTTCAAAATATAGTAAATAATGAAATTAACCTTGTCTATATTTTGGTT 1800
Db 1741 GAAGGATTTCTGTTCAAAATATAGTAAATAATGAAATTAACCTTGTCTATATTTTGGTT 1800
QY 1801 GCAACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCAAGGAGGTGTGTAGT 1860

Db 1801 GCAACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCAAGGAGGTGTGTAGT 1860
QY 1861 TTCTGCTTTTAAATTTCAAGCAAACTGGAATAATTCATCTAATTTATGCTTTCTTTCCC 1920
Db 1861 TTCTGCTTTTAAATTTCAAGCAAACTGGAATAATTCATCTAATTTATGCTTTCTTTCCC 1920
QY 1921 AAGAAGTTTATTAATGATATGCCAGTCTTCTAATTTGGAGACAAAGCTTTAATTTGACAA 1980
Db 1921 AAGAAGTTTATTAATGATATGCCAGTCTTCTAATTTGGAGACAAAGCTTTAATTTGACAA 1980
QY 1981 TGCATTCTAATATATATTTTGTATAGTTACAGTATACAGAGTTGAGTATCCCTTAGATG 2040
Db 1981 TGCATTCTAATATATATTTTGTATAGTTACAGTATACAGAGTTGAGTATCCCTTAGATG 2040
QY 2041 AGATGCTTTGGGACCAAGAGTGTTTTGGATTTTCAATTTATTTTGGATTTTGGATATTTT 2100
Db 2041 AGATGCTTTGGGACCAAGAGTGTTTTGGATTTTCAATTTATTTTGGATTTTGGATATTTT 2100
QY 2101 CCATACATATATGAGAGTTGGAAATGGATTTCAAGTCTAATCATATAAATTCACCTTA 2160
Db 2101 CCATACATATATGAGAGTTGGAAATGGATTTCAAGTCTAATCATATAAATTCACCTTA 2160
QY 2161 TGTGTGATATACACTTATCTGAATAGCTTGAAGTAAATTTTATACATATTTTAAATTA 2220
Db 2161 TGTGTGATATACACTTATCTGAATAGCTTGAAGTAAATTTTATACATATTTTAAATTA 2220
QY 2221 TTTTATGCTGAAACAGAGTTTGGCACAATTTGGACCATCAGAAAGCAGAGTGTCACTAT 2280
Db 2221 TTTTATGCTGAAACAGAGTTTGGCACAATTTGGACCATCAGAAAGCAGAGTGTCACTAT 2280
QY 2281 TTTCAAGTCAGTCTCAAAAAGTTTTCAGATTTAAGTGTGTGATGATGATGATGATGATG 2340
Db 2281 TTTCAAGTCAGTCTCAAAAAGTTTTCAGATTTAAGTGTGTGATGATGATGATGATGATG 2340
QY 2341 TCCGAGTACTTTTGGAGCCAGGAGTGTGATCTTTGAGCCAGGAGTTTGGAGCCAG 2400
Db 2341 TCCGAGTACTTTTGGAGCCAGGAGTGTGATCTTTGAGCCAGGAGTTTGGAGCCAG 2400
QY 2401 ACTGCACAACACAGTGTGAGACCTCTGTTTCTCAAAATAATTTAAATAATTTAGCCAG 2460
Db 2401 ACTGCACAACACAGTGTGAGACCTCTGTTTCTCAAAATAATTTAAATAATTTAGCCAG 2460
QY 2461 GGTGCACACCTGTAGTCCAGGCTACTCAGGAGGTGAGGTAGTGTAGTGTGTGTGAGACTG 2520
Db 2461 GGTGCACACCTGTAGTCCAGGCTACTCAGGAGGTGAGGTAGTGTAGTGTGTGTGAGACTG 2520
QY 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTTGCCACCACTTTCCAGCTTTGGCCACAG 2580
Db 2521 GGAGGTTGAGGCTGAACCTGAGCCAGGATCTTTGCCACCACTTTCCAGCTTTGGCCACAG 2580
QY 2581 TGAGACCTGTCTCAAAAATAAATAAAGTTTTCAGATTTTGGAGCATTTTCAGATCTTCAG 2640
Db 2581 TGAGACCTGTCTCAAAAATAAATAAAGTTTTCAGATTTTGGAGCATTTTCAGATCTTCAG 2640
QY 2641 ATTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCAAGCATTAATCAATAGTGTG 2700
Db 2641 ATTAGGATTTTCAACCTGTACTGACCTTTTGTAGTCAAGCATTAATCAATAGTGTG 2700
QY 2701 GACTCCAGATAACTCAATTTGTGTATACATTTTGTGCTCTCTATTTCAACGAATTTCTAT 2760
Db 2701 GACTCCAGATAACTCAATTTGTGTATACATTTTGTGCTCTCTATTTCAACGAATTTCTAT 2760
QY 2761 GGCCTCTTGTGTGATTTTAAATGCGGAAGGAAACAATAAGAAATTTTGCATTTCTAGA 2820
Db 2761 GGCCTCTTGTGTGATTTTAAATGCGGAAGGAAACAATAAGAAATTTTGCATTTCTAGA 2820
QY 2821 AAAGTCATTTCTGCAAAATATGTCAAGTCTCTGTAGATTTAGCCAAATTTTAGGAAATGAC 2880
Db 2821 AAAGTCATTTCTGCAAAATATGTCAAGTCTCTGTAGATTTAGCCAAATTTTAGGAAATGAC 2880
QY 2881 AAAATTTTACTTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2940

APPLICANT: Hoersht, Sebastian
TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS FOR
IDENTIFICATION, ASSESSMENT, PREVENTION, AND THERAPY
OF CERVICAL CANCER
FILE REFERENCE: MRI-035
CURRENT APPLICATION NUMBER: US/10/171.311
CURRENT FILING DATE: 2002-06-12
PRIOR APPLICATION NUMBER: US 60/298,159
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/298,155
PRIOR FILING DATE: 2001-06-13
PRIOR APPLICATION NUMBER: US 60/335,936
PRIOR FILING DATE: 2001-11-14
NUMBER OF SEQ ID NOS: 238
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 67
LENGTH: 3602
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: 2087, 2093, 2098
OTHER INFORMATION: n = A,T,C or G

US-10-171-311-67

Query Match 1.3%; Score 50; DB 15; Length 3602;
Best Local Similarity 100.0%; Pred. No. 4e-12;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
DB 3501 TAAATGAAAAA 3550

RESULT 5

US-10-001-843-87/c
Sequence 87, Application US/10001843
Publication No. US2002013225A1

GENERAL INFORMATION:
APPLICANT: Salceda, Susana
APPLICANT: Macina, Roberto
APPLICANT: Recipon, Herve
APPLICANT: Caferkey, Robert
APPLICANT: Sun, Yongming
APPLICANT: Liu, Chenghua
APPLICANT: Turner, Leah
TITLE OF INVENTION: Compositions and Methods Relating to Breast Specific Genes and P
FILE REFERENCE: DEX-0267
CURRENT APPLICATION NUMBER: US/10/001.843
CURRENT FILING DATE: 2001-11-20
PRIOR APPLICATION NUMBER: 60/249,992
PRIOR FILING DATE: 2000-11-20
NUMBER OF SEQ ID NOS: 218
SOFTWARE: PatentIn version 3.1
SEQ ID NO 87
LENGTH: 476
TYPE: DNA
ORGANISM: Homo sapien

US-10-001-843-87

Query Match 1.3%; Score 49; DB 14; Length 476;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 126 AAATGAAAAA 78

RESULT 6

US-10-424-599-35023
Sequence 35023, Application US/10424599
Publication No. US20040031072A1

GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei

TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 35023
LENGTH: 606
TYPE: DNA
ORGANISM: Glycine max

NAME/KEY: unsure
LOCATION: (1)-(606)
OTHER INFORMATION: unsure at all n locations
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MFT3847_131628C.1
US-10-424-599-35023

Query Match 1.3%; Score 49; DB 13; Length 606;
Best Local Similarity 100.0%; Pred. No. 1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 85 AAATGAAAAA 133

RESULT 7

US-10-106-698-741
Sequence 741, Application US/10106698
Publication No. US20030109690A1

GENERAL INFORMATION:
APPLICANT: Ruben et al.
TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptides
FILE REFERENCE: PA005PI
CURRENT APPLICATION NUMBER: US/10/106,698
CURRENT FILING DATE: 2002-03-27
PRIOR APPLICATION NUMBER: PCT/US00/26524
PRIOR FILING DATE: 2000-09-28
PRIOR APPLICATION NUMBER: US 60/157,137
PRIOR FILING DATE: 1999-09-29
PRIOR APPLICATION NUMBER: US 60/163,280
PRIOR FILING DATE: 1999-11-03
NUMBER OF SEQ ID NOS: 8564
SOFTWARE: PatentIn Ver. 3.0
SEQ ID NO 741
LENGTH: 1057
TYPE: DNA
ORGANISM: Homo sapiens
NAME/KEY: misc feature
LOCATION: (1056)-(1057)
OTHER INFORMATION: n equals a,t,g, or c

US-10-106-698-741

Query Match 1.3%; Score 49; DB 15; Length 1057;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 1006 AAATGAAAAA 1054

RESULT 8

US-09-925-301-10
Sequence 10, Application US/09925301
Patent No. US20020052308A1


```
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 10
; LENGTH: 1373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1364)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (1373)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-10

Query Match      1.3%; Score 49; DB 9; Length 1373;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
Db 1294 AAATGAAAAA 1342

RESULT 9
US-10-198-846-9741/c
; Sequence 9741, Application US/10198846
; Publication No. US20030099741
; GENERAL INFORMATION:
; APPLICANT: Lillie, James
; APPLICANT: Xu, Yongyao
; APPLICANT: Wang, Youzhen
; APPLICANT: Steinmann, Kathleen
; TITLE OF INVENTION: NOVEL GENES, COMPOSITIONS, KITS, AND METHODS
; TITLE OF INVENTION: FOR IDENTIFICATION, ASSESSMENT, PREVENTION, AND
; TITLE OF INVENTION: THERAPY OF BREAST CANCER
; FILE REFERENCE: MRI-049
; CURRENT APPLICATION NUMBER: US/10/198,846
; CURRENT FILING DATE: 2002-07-18
; PRIOR APPLICATION NUMBER: 60/306,220
; PRIOR FILING DATE: 2001-07-18
; NUMBER OF SEQ ID NOS: 14084
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9741
; LENGTH: 1392
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-198-846-9741

Query Match      1.3%; Score 49; DB 15; Length 1392;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
Db 374 AAATGAAAAA 326

RESULT 10
US-10-424-599-11755
; Sequence 11755, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 11755
; LENGTH: 1417
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_110622C.1
US-10-424-599-11755

Query Match      1.3%; Score 49; DB 13; Length 1417;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
Db 1354 AAATGAAAAA 1402

RESULT 11
US-10-424-599-106798
; Sequence 106798, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 106798
; LENGTH: 1423
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_67454C.1
US-10-424-599-106798

Query Match      1.3%; Score 49; DB 13; Length 1423;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
Db 1339 AAATGAAAAA 1387

RESULT 12
US-10-007-926A-34
; Sequence 34, Application US/10007926A
; Publication No. US20030143539A1
; GENERAL INFORMATION:
; APPLICANT: BERTUCCI, FRANCOIS
; APPLICANT: HOULGATTE, REMI
; APPLICANT: BIRNBAUM, DANIEL
; APPLICANT: NGUYEN, CATHERINE
; APPLICANT: VIENS, PATRICE
; APPLICANT: FERT, VINCENT
; TITLE OF INVENTION: GENE EXPRESSION PROFILING OF PRIMARY BREAST CARCINOMAS
; TITLE OF INVENTION: USING ARRAYS OF CANDIDATE GENES
; FILE REFERENCE: 1546-R-00
```

	/	CURRENT APPLICATION NUMBER:	US/10/007,926A	/	PRIOR FILING DATE:	2000-08-14
	/	CURRENT FILING DATE:	2001-12-07	/	PRIOR APPLICATION NUMBER:	60/220,963
	/	PRIOR APPLICATION NUMBER:	60/254,090	/	PRIOR FILING DATE:	2000-07-26
	/	PRIOR FILING DATE:	2000-12-08	/	PRIOR APPLICATION NUMBER:	60/217,496
	/	NUMBER OF SEQ ID NOS:	468	/	PRIOR FILING DATE:	2000-07-11
	/	SOFTWARE:	PatentIn Ver. 2.1	/	PRIOR APPLICATION NUMBER:	60/225,447
	/	SEQ ID NO 34		/	PRIOR FILING DATE:	2000-08-14
	/	LENGTH:	1429	/	PRIOR APPLICATION NUMBER:	60/218,290
	/	TYPE:	DNA	/	PRIOR FILING DATE:	2000-07-14
	/	ORGANISM:	Homo sapiens	/	PRIOR APPLICATION NUMBER:	60/225,757
	/	FEATURE:		/	PRIOR FILING DATE:	2000-08-14
	/	OTHER INFORMATION:	protein phosphatase 4 (formerly x),	/	PRIOR APPLICATION NUMBER:	60/226,868
	/	OTHER INFORMATION:	catalytic subunit (ppp4C) gene.	/	PRIOR FILING DATE:	2000-08-22
	/	US-10-007-926A-34		/	PRIOR APPLICATION NUMBER:	60/216,647
	/			/	PRIOR FILING DATE:	2000-07-07
	/	Query Match	1.3%; Score 49; DB 15; Length 1429;	/	PRIOR APPLICATION NUMBER:	60/225,267
	/	Best Local Similarity	100.0%; Pred. No. 1.le-11;	/	PRIOR FILING DATE:	2000-08-14
	/	Matches	49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/	PRIOR APPLICATION NUMBER:	60/216,880
	/			/	PRIOR FILING DATE:	2000-07-07
	/	Qy	3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762	/	PRIOR APPLICATION NUMBER:	60/225,270
	/			/	PRIOR FILING DATE:	2000-08-14
	/	Db	1336 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1384	/	PRIOR APPLICATION NUMBER:	60/251,869
	/			/	PRIOR FILING DATE:	2000-12-08
	/			/	PRIOR APPLICATION NUMBER:	60/235,834
	/	RESULT 13		/	PRIOR FILING DATE:	2000-09-27
	/	US-09-764-903-19		/	PRIOR APPLICATION NUMBER:	60/234,274
	/	Sequence 19, Application US/09764903		/	PRIOR FILING DATE:	2000-09-21
	/	Patent No. US20020090674A1		/	PRIOR APPLICATION NUMBER:	60/234,223
	/	GENERAL INFORMATION:		/	PRIOR FILING DATE:	2000-09-21
	/	APPLICANT:	Rosen et al.	/	PRIOR APPLICATION NUMBER:	60/228,924
	/	TITLE OF INVENTION:	Nucleic Acids, Proteins, and Antibodies	/	PRIOR FILING DATE:	2000-08-30
	/	FILE REFERENCE:	PTZ28	/	PRIOR APPLICATION NUMBER:	60/224,518
	/	CURRENT APPLICATION NUMBER:	US/09/764,903	/	PRIOR FILING DATE:	2000-08-14
	/	CURRENT FILING DATE:	2001-01-17	/	PRIOR APPLICATION NUMBER:	60/236,369
	/	Prior application data removed - consult PALM or file wrapper		/	PRIOR FILING DATE:	2000-09-29
	/	NUMBER OF SEQ ID NOS:	67	/	PRIOR APPLICATION NUMBER:	60/224,519
	/	SOFTWARE:	PatentIn Ver. 2.0	/	PRIOR FILING DATE:	2000-08-14
	/	SEQ ID NO 19		/	PRIOR APPLICATION NUMBER:	60/220,964
	/	LENGTH:	1998	/	PRIOR FILING DATE:	2000-07-26
	/	TYPE:	DNA	/	PRIOR APPLICATION NUMBER:	60/241,809
	/	ORGANISM:	Homo sapiens	/	PRIOR FILING DATE:	2000-10-20
	/	US-09-764-903-19		/	PRIOR APPLICATION NUMBER:	60/249,299
	/			/	PRIOR FILING DATE:	2000-11-17
	/	Query Match	1.3%; Score 49; DB 9; Length 1998;	/	PRIOR APPLICATION NUMBER:	60/236,327
	/	Best Local Similarity	100.0%; Pred. No. 1.le-11;	/	PRIOR FILING DATE:	2000-09-29
	/	Matches	49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	/	PRIOR APPLICATION NUMBER:	60/241,785
	/			/	PRIOR FILING DATE:	2000-10-20
	/	Qy	3714 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762	/	PRIOR APPLICATION NUMBER:	60/244,617
	/			/	PRIOR FILING DATE:	2000-11-01
	/	Db	1906 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1954	/	PRIOR APPLICATION NUMBER:	60/225,268
	/			/	PRIOR FILING DATE:	2000-08-14
	/			/	PRIOR APPLICATION NUMBER:	60/236,368
	/	RESULT 14		/	PRIOR FILING DATE:	2000-09-29
	/	US-10-091-391-19		/	PRIOR APPLICATION NUMBER:	60/251,856
	/	Sequence 19, Application US/10091391		/	PRIOR FILING DATE:	2000-12-08
	/	Publication No. US20030082691A1		/	PRIOR APPLICATION NUMBER:	60/251,868
	/	GENERAL INFORMATION:		/	PRIOR FILING DATE:	2000-12-08
	/	APPLICANT:	Rosen et al.	/	PRIOR APPLICATION NUMBER:	60/229,344
	/	TITLE OF INVENTION:	Nucleic Acids, Proteins, and Antibodies	/	PRIOR FILING DATE:	2000-09-01
	/	FILE REFERENCE:	FTZ28C1	/	PRIOR APPLICATION NUMBER:	60/234,997
	/	CURRENT APPLICATION NUMBER:	US/10/091,391	/	PRIOR FILING DATE:	2000-09-25
	/	CURRENT FILING DATE:	2002-03-07	/	PRIOR APPLICATION NUMBER:	60/229,343
	/	PRIOR APPLICATION NUMBER:	09/764,903	/	PRIOR FILING DATE:	2000-09-01
	/	PRIOR FILING DATE:	2001-01-17	/	PRIOR APPLICATION NUMBER:	60/229,345
	/	PRIOR APPLICATION NUMBER:	60/179,065	/	PRIOR FILING DATE:	2000-09-01
	/	PRIOR FILING DATE:	20			

; PRIOR APPLICATION NUMBER: 60/236,367
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/237,039
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,038
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/236,370
; PRIOR FILING DATE: 2000-09-29
; PRIOR APPLICATION NUMBER: 60/236,802
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,037
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/237,040
; PRIOR FILING DATE: 2000-10-02
; PRIOR APPLICATION NUMBER: 60/240,960
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/239,935
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/239,937
; PRIOR FILING DATE: 2000-10-13
; PRIOR APPLICATION NUMBER: 60/241,787
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,474
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/246,532
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/249,216
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,210
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/226,681
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,759
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/225,213
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/227,182
; PRIOR FILING DATE: 2000-08-22
; PRIOR APPLICATION NUMBER: 60/225,214
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/235,836
; PRIOR FILING DATE: 2000-09-27
; PRIOR APPLICATION NUMBER: 60/230,438
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/215,135
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: 60/225,266
; PRIOR FILING DATE: 2000-08-14
; PRIOR APPLICATION NUMBER: 60/249,218
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,208
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,213
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,212
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,207
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,245
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,244
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,217
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,211
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,215
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,264
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,214
; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/249,297

; PRIOR FILING DATE: 2000-11-17
; PRIOR APPLICATION NUMBER: 60/232,400
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/231,242
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,081
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/232,080
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,414
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/231,244
; PRIOR FILING DATE: 2000-09-08
; PRIOR APPLICATION NUMBER: 60/233,064
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/233,063
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,397
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,399
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/232,401
; PRIOR FILING DATE: 2000-09-14
; PRIOR APPLICATION NUMBER: 60/241,808
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,826
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,786
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,221
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/246,475
; PRIOR FILING DATE: 2000-11-08
; PRIOR APPLICATION NUMBER: 60/231,243
; PRIOR FILING DATE: 2000-09-08

Query Match 1.3% Score 49; DB 15; Length 1998;

Best Local Similarity 100.0%; Pred. No. 1.1e-11; Mismatches 0; Indels 0; Gaps 0;

Qy 3714 AAATGAAAAA
Db 1906 AAATGAAAAA

RESULT 15

US-10-264-049-740/c
; Sequence 740, Application US/10264049
; Publication No. US20040005579A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA133P1
; CURRENT APPLICATION NUMBER: US/10/264,049
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/18569
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/209,467
; PRIOR FILING DATE: 2000-06-07
; NUMBER OF SEQ ID NOS: 4360
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 740
; LENGTH: 2143
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)..(1)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (10)..(10)
; OTHER INFORMATION: n equals a,t,g, or c

Db 2650 AAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2698

APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 34
LENGTH: 2713
TYPE: DNA
ORGANISM: Homo sapiens
US-10-436-523-34

Query Match 1.3%; Score 49; DB 15; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 2650 AAATGAAAAA... 2698

RESULT 21

US-10-128-714-1558/c
Sequence 1558, Application US/10128714
Publication No. US20030119013A1

GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wengqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M
APPLICANT: Lemieux, Sebastien M

TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
METHODS OF USE

FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 1558
LENGTH: 2944
TYPE: DNA
ORGANISM: Aspergillus fumigatus
US-10-128-714-1558

Query Match 1.3%; Score 49; DB 15; Length 2944;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 941 AAATGAAAAA... 893

RESULT 22

US-10-128-714-6558/c
Sequence 6558, Application US/10128714
Publication No. US20030119013A1

GENERAL INFORMATION:
APPLICANT: Jiang, Bo
APPLICANT: Hu, Wengqi
APPLICANT: Tishkoff, Daniel
APPLICANT: Zamudio, Carlos
APPLICANT: Eroshkin, Alexey M

APPLICANT: Lemieux, Sebastien M
TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
FILE REFERENCE: 10182-018-999
CURRENT APPLICATION NUMBER: US/10/128,714
CURRENT FILING DATE: 2002-04-23
PRIOR APPLICATION NUMBER: US 60/285,697
PRIOR FILING DATE: 2001-04-23
PRIOR APPLICATION NUMBER: US 60/287,066
PRIOR FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/295,890
PRIOR FILING DATE: 2001-06-05
PRIOR APPLICATION NUMBER: US 60/303,899
PRIOR FILING DATE: 2001-07-09
PRIOR APPLICATION NUMBER: US 60/316,362
PRIOR FILING DATE: 2001-08-31
NUMBER OF SEQ ID NOS: 8603
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6558
LENGTH: 3362
TYPE: DNA
ORGANISM: Aspergillus fumigatus
US-10-128-714-6558

Query Match 1.3%; Score 49; DB 15; Length 3362;
Best Local Similarity 100.0%; Pred. No. 1.1e-11;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 1359 AAATGAAAAA... 1311

RESULT 23

US-09-745-763-20
Sequence 20, Application US/09745763
Patent No. US20020065394A1

GENERAL INFORMATION:

APPLICANT: Jacobs, Kenneth
McCoy, John M.
Lavalie, Edward R.
Collins-Racie, Lisa A.
Evans, Cheryl
Merberg, David
Treacy, Maurice
Spaulding, Vikki

TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
ENCODING THEM

NUMBER OF SEQUENCES: 219

CORRESPONDENCE ADDRESS:

ADDRESSEE: Genetics Institute, Inc.
STREET: 87 CambridgePark Drive
CITY: Cambridge
STATE: MA
COUNTRY: U.S.A.
ZIP: 02140

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/745,763
FILING DATE: 18-Jun-2000
CLASSIFICATION: <Unknown>

ATTORNEY/AGENT INFORMATION:

NAME: Sprunger, Suzanne A.
REGISTRATION NUMBER: 41,323

TELECOMMUNICATION INFORMATION:

TELEPHONE: (617) 498-8284

TELEFAX: (617) 876-5851

INFORMATION FOR SEQ ID NO: 20:

SEQUENCE CHARACTERISTICS:

```
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; MOLECULE TYPE: CDNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 20:
US-09-745-763-20

Query Match
Best Local Similarity 1.3%; Score 49; DB 9; Length 4237;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA... 3762
Db 2422 AATGAAAAA... 2470

RESULT 24
US-10-128-714-558/c
; Sequence 558, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Ershkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 558
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-558

Query Match
Best Local Similarity 1.3%; Score 49; DB 15; Length 4944;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA... 3762
Db 1941 AATGAAAAA... 1893

RESULT 25
US-10-128-714-5558/c
; Sequence 5558, Application US/10128714
; Publication No. US20030119013A1
; GENERAL INFORMATION:
; APPLICANT: Jiang, Bo
; APPLICANT: Hu, Wenqi
; APPLICANT: Tishkoff, Daniel
; APPLICANT: Zamudio, Carlos
; APPLICANT: Ershkin, Alexey M
; APPLICANT: Lemieux, Sebastien M
; TITLE OF INVENTION: Identification of Essential Genes in Aspergillus fumigatus and
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 558
; LENGTH: 4944
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-558

Query Match
Best Local Similarity 1.3%; Score 49; DB 15; Length 4944;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA... 3762
Db 1941 AATGAAAAA... 1893

RESULT 26
US-09-960-352-6036
; Sequence 6036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6036
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3058-015-Q1-X1-G5
US-09-960-352-6036

Query Match
Best Local Similarity 1.3%; Score 48; DB 9; Length 337;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA... 3762
Db 138 AATGAAAAA... 185

RESULT 27
US-10-085-783A-16208
; Sequence 16208, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
```

```
; FILE REFERENCE: 10182-018-999
; CURRENT APPLICATION NUMBER: US/10/128,714
; CURRENT FILING DATE: 2002-04-23
; PRIOR APPLICATION NUMBER: US 60/285,697
; PRIOR FILING DATE: 2001-04-23
; PRIOR APPLICATION NUMBER: US 60/287,066
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/295,890
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/303,899
; PRIOR FILING DATE: 2001-07-09
; PRIOR APPLICATION NUMBER: US 60/316,362
; PRIOR FILING DATE: 2001-08-31
; NUMBER OF SEQ ID NOS: 8603
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 558
; LENGTH: 5362
; TYPE: DNA
; ORGANISM: Aspergillus fumigatus
US-10-128-714-5558

Query Match
Best Local Similarity 1.3%; Score 49; DB 15; Length 5362;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA... 3762
Db 2359 AATGAAAAA... 2311

RESULT 26
US-09-960-352-6036
; Sequence 6036, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 6036
; LENGTH: 337
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 26-LIB3058-015-Q1-X1-G5
US-09-960-352-6036

Query Match
Best Local Similarity 1.3%; Score 48; DB 9; Length 337;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA... 3762
Db 138 AATGAAAAA... 185

RESULT 27
US-10-085-783A-16208
; Sequence 16208, Application US/10085783A
; Publication No. US20040037841A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2002
; CURRENT APPLICATION NUMBER: US/10/085,783A
; CURRENT FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
```

```
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 16208
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Human
US-10-085-783A-16208

Query Match
Best Local Similarity 1.3%; Score 48; DB 13; Length 348;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762
DB 271 AATGAAAAA 318

RESULT 28
US-10-242-535A-16208
; Sequence 16208, Application US/10242535A
; Publication No. US20040013663A1
; GENERAL INFORMATION:
; APPLICANT: ChondroGene Inc.
; APPLICANT: Liaw, C.C.
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis
; FILE REFERENCE: 4231/2005
; CURRENT APPLICATION NUMBER: US/10/242.535A
; CURRENT FILING DATE: 2002-09-12
; PRIOR APPLICATION NUMBER: US 10/085,783
; PRIOR FILING DATE: 2002-02-28
; PRIOR APPLICATION NUMBER: US 60/305,340
; PRIOR FILING DATE: 2001-07-13
; PRIOR APPLICATION NUMBER: US 60/275,017
; PRIOR FILING DATE: 2001-03-12
; PRIOR APPLICATION NUMBER: US 60/271,955
; PRIOR FILING DATE: 2001-02-28
; NUMBER OF SEQ ID NOS: 58994
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 16208
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Human
US-10-242-535A-16208

Query Match
Best Local Similarity 1.3%; Score 48; DB 16; Length 348;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762
DB 271 AATGAAAAA 318

RESULT 29
US-10-424-599-47139
; Sequence 47139, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 16208
; LENGTH: 348
; TYPE: DNA
; ORGANISM: Human
US-10-424-599-47139

Query Match
Best Local Similarity 1.3%; Score 48; DB 13; Length 372;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762
DB 58 AATGAAAAA 105

RESULT 30
US-09-918-995-16319
; Sequence 16319, Application US/09918995
; Publication No. US20030073623A1
; GENERAL INFORMATION:
; APPLICANT: Hysed, Inc.
; TITLE OF INVENTION: NOVEL NUCLEIC ACID SEQUENCES OBTAINED
; FILE REFERENCE: 20411-756
; CURRENT APPLICATION NUMBER: US/09/918.995
; CURRENT FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: US/09/235,076
; PRIOR FILING DATE: 1999-01-20
; NUMBER OF SEQ ID NOS: 38054
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 16319
; LENGTH: 401
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-918-995-16319

Query Match
Best Local Similarity 1.3%; Score 48; DB 10; Length 401;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762
DB 84 AATGAAAAA 131

RESULT 31
US-10-424-599-54555
; Sequence 54555, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424.599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 54555
; LENGTH: 462
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20276C.1
US-10-424-599-54555

Query Match
Best Local Similarity 1.3%; Score 48; DB 13; Length 462;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 120 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 32

US-10-424-599-46783/c
; Sequence 46783, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 46783
; LENGTH: 544
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE: Glycine max
; OTHER INFORMATION: Clone ID: PAT_MRT3847_142247C.1
US-10-424-599-46783

Query Match 1.3%; Score 48; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 57 AATGAAAAA 1.3%; Score 48; DB 13; Length 544;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 33

US-10-351-334-58
; Sequence 58, Application US/10351334
; Publication No. US20040034196A1
; GENERAL INFORMATION:
; APPLICANT: Komatsoulis et al
; TITLE OF INVENTION: 98 Human Secreted Proteins
; FILE REFERENCE: P2031P2
; CURRENT APPLICATION NUMBER: US/10/351,334
; CURRENT FILING DATE: 2003-01-27
; PRIOR APPLICATION NUMBER: 60/350,898
; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: 09/489,847
; PRIOR FILING DATE: 2000-01-24
; PRIOR APPLICATION NUMBER: PCT/US99/17130
; PRIOR FILING DATE: 1999-07-29
; PRIOR APPLICATION NUMBER: 60/094,657
; PRIOR FILING DATE: 1998-07-30
; PRIOR APPLICATION NUMBER: 60/095,486
; PRIOR FILING DATE: 1998-08-05
; PRIOR APPLICATION NUMBER: 60/096,319
; PRIOR FILING DATE: 1998-08-12
; PRIOR APPLICATION NUMBER: 60/095,454
; PRIOR FILING DATE: 1998-08-06
; PRIOR APPLICATION NUMBER: 60/095,455
; PRIOR FILING DATE: 1998-08-06
; NUMBER OF SEQ ID NOS: 376
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 58
; LENGTH: 619
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (526)

; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (619)
; OTHER INFORMATION: n equals a,t,g, or c
US-10-351-334-58

Query Match 1.3%; Score 48; DB 13; Length 619;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 619;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 566 AATGAAAAA 1.3%; Score 48; DB 13; Length 619;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 34

US-10-424-599-95989
; Sequence 95989, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95989
; LENGTH: 751
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)-(751)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_5768C.1
US-10-424-599-95989

Query Match 1.3%; Score 48; DB 13; Length 751;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAA 1.3%; Score 48; DB 13; Length 751;
Best Local Similarity 100.0%; Pred. No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Db 513 AATGAAAAA 1.3%; Score 48; DB 13; Length 751;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 35

US-09-770-445-579/c
; Sequence 579, Application US/09770445
; Patent No. US20020023281A1
; GENERAL INFORMATION:
; APPLICANT: Gorlach, Jorn
; APPLICANT: An, Yong-Oiang
; APPLICANT: Hamilton, Carol M.
; APPLICANT: Price, Jennifer L.
; APPLICANT: Raines, Tracy M.
; APPLICANT: Yu, Yang
; APPLICANT: Rameaka, Joshua G.
; APPLICANT: Page, Amy
; APPLICANT: Matthew, Abraham V.
; APPLICANT: Ledford, Brooke L.
; APPLICANT: Woessner, Jeffrey P.
; APPLICANT: Haas, William David
; APPLICANT: Garcia, Carlos A.
; APPLICANT: Krickler, Naja
; APPLICANT: Slader, Ted
; APPLICANT: Davis, Keith R.


```
; APPLICANT: Allen, Keith
; APPLICANT: Hoffman, Neil
; APPLICANT: Hubban, Patrick
; TITLE OF INVENTION: Expressed Sequences of Arabidopsis
; FILE REFERENCE: 2023US (PAPA-012PRV)
; CURRENT APPLICATION NUMBER: US/09/770,445
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: US 60/178,472
; PRIOR FILING DATE: 2000-01-27
; NUMBER OF SEQ ID NOS: 999
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 579
; LENGTH: 865
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-770-445-579

Query Match      1.3%; Score 48; DB 9; Length 865;
Best Local Similarity 100.0%; Pred.No. 3e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 54 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 7

RESULT 36
US-10-219-535-39
; Sequence 39, Application US/10219535
; Publication No. US20040044179A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC60
; CURRENT APPLICATION NUMBER: US/10/219,535
; CURRENT FILING DATE: 2002-08-14
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
```

```
; ORGANISM: Homo Sapien
US-10-219-535-39

Query Match      1.3%; Score 48; DB 13; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 37
US-10-232-230-39
; Sequence 39, Application US/10232230
; Publication No. US20040044180A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC103
; CURRENT APPLICATION NUMBER: US/10/232,230
; CURRENT FILING DATE: 2002-08-29
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-232-230-39

Query Match      1.3%; Score 48; DB 13; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 38
US-10-232-224-39
; Sequence 39, Application US/10232224
```

```
Publication No. US20030065147A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3530P1C179
CURRENT APPLICATION NUMBER: US/10/227,884
CURRENT FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
```

```
Publication No. US20030065147A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3530P1C111
CURRENT APPLICATION NUMBER: US/10/232,224
CURRENT FILING DATE: 2002-08-29
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 246
SEQ ID NO 39
LENGTH: 1837
TYPE: DNA
ORGANISM: Homo Sapien
US-10-232-224-39
Query Match 1.3%; Score 48; DB 13; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Caps 0;
Qy 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
RESULT 39
US-10-227-884-39
Sequence 39, Application US/10227884
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
TITLE OF INVENTION: ACIDS ENCODING THE SAME
```

; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022

; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.2e-11;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 40

US-10-230-163-39
; Sequence 39, Application US/10230163
; Publication No. US20030036635A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watarabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530PIC96
; CURRENT APPLICATION NUMBER: US/10/230,163
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17

, PRIOR APPLICATION NUMBER: 60/062287
, PRIOR FILING DATE: 1997-10-17
, PRIOR APPLICATION NUMBER: 60/063549
, PRIOR FILING DATE: 1997-10-28
, PRIOR APPLICATION NUMBER: 60/064103
, PRIOR FILING DATE: 1997-10-31
, PRIOR APPLICATION NUMBER: 60/069873
, PRIOR FILING DATE: 1997-12-17
, PRIOR APPLICATION NUMBER: 60/078910
, PRIOR FILING DATE: 1998-03-20
, PRIOR APPLICATION NUMBER: 60/079294
, PRIOR FILING DATE: 1998-03-25
, PRIOR APPLICATION NUMBER: 60/079656
, PRIOR FILING DATE: 1998-03-26
, PRIOR APPLICATION NUMBER: 60/079728
, PRIOR FILING DATE: 1998-03-27
, PRIOR APPLICATION NUMBER: 60/081819
, PRIOR FILING DATE: 1998-04-15
, PRIOR APPLICATION NUMBER: 60/081955
, PRIOR FILING DATE: 1998-04-15
, PRIOR APPLICATION NUMBER: 60/082804
, PRIOR FILING DATE: 1998-04-22
, PRIOR APPLICATION NUMBER: 60/084441
, PRIOR FILING DATE: 1998-05-06
, PRIOR APPLICATION NUMBER: 60/085323
, PRIOR FILING DATE: 1998-05-13
, PRIOR APPLICATION NUMBER: 60/085579
, PRIOR FILING DATE: 1998-05-15
, PRIOR APPLICATION NUMBER: 60/086392
, PRIOR FILING DATE: 1998-05-22
, PRIOR APPLICATION NUMBER: 60/089532
, PRIOR FILING DATE: 1998-06-17
, PRIOR APPLICATION NUMBER: 60/089538
, PRIOR FILING DATE: 1998-06-17
, PRIOR APPLICATION NUMBER: 60/089905
, PRIOR FILING DATE: 1998-06-18
, PRIOR APPLICATION NUMBER: 60/090472
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090557
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090691
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090695
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/091982
, PRIOR FILING DATE: 1998-07-07
, PRIOR APPLICATION NUMBER: 60/095302
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095318
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095916
, PRIOR FILING DATE: 1998-08-10
, PRIOR APPLICATION NUMBER: 60/096146
, PRIOR FILING DATE: 1998-08-11
, PRIOR APPLICATION NUMBER: 60/096791
, PRIOR FILING DATE: 1998-08-17
, PRIOR APPLICATION NUMBER: 60/097986
, PRIOR FILING DATE: 1998-08-26
, PRIOR APPLICATION NUMBER: 60/098544
, PRIOR FILING DATE: 1998-08-31
, PRIOR APPLICATION NUMBER: 60/099596
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099598
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099803
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099811
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099812
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099816
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/100038
, PRIOR FILING DATE: 1998-09-11
, PRIOR APPLICATION NUMBER: 60/100385
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100390
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100627
, PRIOR FILING DATE: 1998-09-16
, PRIOR APPLICATION NUMBER: 60/100848
, PRIOR FILING DATE: 1998-09-18
, PRIOR APPLICATION NUMBER: 60/100919
, PRIOR FILING DATE: 1998-09-17
, PRIOR APPLICATION NUMBER: 60/101477
, PRIOR FILING DATE: 1998-09-23
, PRIOR APPLICATION NUMBER: 60/101738
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101741
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101786
, PRIOR FILING DATE: 1998-09-25
, PRIOR APPLICATION NUMBER: 60/101916
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101922
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/106178
, PRIOR FILING DATE: 1998-10-28
, PRIOR APPLICATION NUMBER: 60/106248
, PRIOR FILING DATE: 1998-10-29
, PRIOR APPLICATION NUMBER: 60/106464
, PRIOR FILING DATE: 1998-10-30
, PRIOR APPLICATION NUMBER: 60/106905
, PRIOR FILING DATE: 1998-11-03
, PRIOR APPLICATION NUMBER: 60/108787
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108801
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108849
, PRIOR FILING DATE: 1998-11-18
, PRIOR APPLICATION NUMBER: 60/112422
, PRIOR FILING DATE: 1998-12-15
, PRIOR APPLICATION NUMBER: 60/113296
, PRIOR FILING DATE: 1998-12-22
, PRIOR APPLICATION NUMBER: 60/113605
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/113621
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/115558
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115565
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115733
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/119549
, PRIOR FILING DATE: 1999-02-10
, PRIOR APPLICATION NUMBER: 60/123618
, PRIOR FILING DATE: 1999-03-10
, PRIOR APPLICATION NUMBER: 60/125259
, PRIOR FILING DATE: 1999-03-19
, PRIOR APPLICATION NUMBER: 60/125775
, PRIOR FILING DATE: 1999-03-23
, PRIOR APPLICATION NUMBER: 60/126773
, PRIOR FILING DATE: 1999-03-29
, PRIOR APPLICATION NUMBER: 60/127887
, PRIOR FILING DATE: 1999-04-05
, PRIOR APPLICATION NUMBER: 60/130232
, PRIOR FILING DATE: 1999-04-21
, PRIOR APPLICATION NUMBER: 60/131022
, PRIOR FILING DATE: 1999-04-26
, PRIOR APPLICATION NUMBER: 60/131270
, PRIOR FILING DATE: 1999-04-27
, PRIOR APPLICATION NUMBER: 60/131291
, PRIOR FILING DATE: 1999-04-27
, PRIOR APPLICATION NUMBER: 60/131445
, PRIOR FILING DATE: 1999-04-28

```
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/145963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835
```

```
Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3715 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
```

RESULT 41

```
US-10-230-338-39
; Sequence 39, Application US/10230338
; Publication No. US2003004934A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC92
; CURRENT APPLICATION NUMBER: US/10/230,338
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
```

```
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-338-39
```

```
Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

```
QY 3715 ATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
```

RESULT 42

```
US-10-218-631-39
; Sequence 39, Application US/10218631
; Publication No. US20030045897A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC14
; CURRENT APPLICATION NUMBER: US/10/218,631
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
```


PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21
PRIOR APPLICATION NUMBER: 60/131022
PRIOR FILING DATE: 1999-04-26
PRIOR APPLICATION NUMBER: 60/131270
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131291
PRIOR FILING DATE: 1999-04-27
PRIOR APPLICATION NUMBER: 60/131445
PRIOR FILING DATE: 1999-04-28
PRIOR APPLICATION NUMBER: 60/134287
PRIOR FILING DATE: 1999-05-14
PRIOR APPLICATION NUMBER: 60/140650
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/140723
PRIOR FILING DATE: 1999-06-22
PRIOR APPLICATION NUMBER: 60/141037
PRIOR FILING DATE: 1999-06-23
PRIOR APPLICATION NUMBER: 60/144758
PRIOR FILING DATE: 1999-07-20
PRIOR APPLICATION NUMBER: 60/145698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: 60/146222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: 60/146963
PRIOR FILING DATE: 1999-08-03
PRIOR APPLICATION NUMBER: 60/149320
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/149638
PRIOR FILING DATE: 1999-08-17
PRIOR APPLICATION NUMBER: 60/151733
PRIOR FILING DATE: 1999-08-31
PRIOR APPLICATION NUMBER: 60/164418
PRIOR FILING DATE: 1999-11-09
PRIOR APPLICATION NUMBER: 60/166361
PRIOR FILING DATE: 1999-11-16
PRIOR APPLICATION NUMBER: 60/169445
PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169495

PRIOR FILING DATE: 1999-12-07
PRIOR APPLICATION NUMBER: 60/169835
Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAA 3762
DB 1729 AATGAA 1776
RESULT 47
US-10-227-883-39
Sequence 39, Application US/10227883
Publication No. US20030073817A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe P.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530PIC78
CURRENT APPLICATION NUMBER: US/10/227,883
CURRENT FILING DATE: 2002-08-26
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062387
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18

, PRIOR APPLICATION NUMBER: 60/090472
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090557
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090691
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090695
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/091982
, PRIOR FILING DATE: 1998-07-07
, PRIOR APPLICATION NUMBER: 60/095302
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095318
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095916
, PRIOR FILING DATE: 1998-08-10
, PRIOR APPLICATION NUMBER: 60/096146
, PRIOR FILING DATE: 1998-08-11
, PRIOR APPLICATION NUMBER: 60/096791
, PRIOR FILING DATE: 1998-08-17
, PRIOR APPLICATION NUMBER: 60/097986
, PRIOR FILING DATE: 1998-08-26
, PRIOR APPLICATION NUMBER: 60/098544
, PRIOR FILING DATE: 1998-08-31
, PRIOR APPLICATION NUMBER: 60/099596
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099598
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099803
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099811
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099812
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099816
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/100038
, PRIOR FILING DATE: 1998-09-11
, PRIOR APPLICATION NUMBER: 60/100385
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100390
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100627
, PRIOR FILING DATE: 1998-09-16
, PRIOR APPLICATION NUMBER: 60/100848
, PRIOR FILING DATE: 1998-09-18
, PRIOR APPLICATION NUMBER: 60/100919
, PRIOR FILING DATE: 1998-09-17
, PRIOR APPLICATION NUMBER: 60/101477
, PRIOR FILING DATE: 1998-09-23
, PRIOR APPLICATION NUMBER: 60/101738
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101741
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101786
, PRIOR FILING DATE: 1998-09-25
, PRIOR APPLICATION NUMBER: 60/101916
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101922
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/106178
, PRIOR FILING DATE: 1998-10-28
, PRIOR APPLICATION NUMBER: 60/106248
, PRIOR FILING DATE: 1998-10-29
, PRIOR APPLICATION NUMBER: 60/106464
, PRIOR FILING DATE: 1998-10-30
, PRIOR APPLICATION NUMBER: 60/106905
, PRIOR FILING DATE: 1998-11-03
, PRIOR APPLICATION NUMBER: 60/108787
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108801
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108849

, PRIOR FILING DATE: 1998-11-18
, PRIOR APPLICATION NUMBER: 60/112422
, PRIOR FILING DATE: 1998-12-15
, PRIOR APPLICATION NUMBER: 60/113296
, PRIOR FILING DATE: 1998-12-22
, PRIOR APPLICATION NUMBER: 60/113605
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/113621
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/115558
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115565
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115733
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/119549
, PRIOR FILING DATE: 1999-02-10
, PRIOR APPLICATION NUMBER: 60/123618
, PRIOR FILING DATE: 1999-03-10
, PRIOR APPLICATION NUMBER: 60/125259
, PRIOR FILING DATE: 1999-03-19
, PRIOR APPLICATION NUMBER: 60/125775
, PRIOR FILING DATE: 1999-03-23
, PRIOR APPLICATION NUMBER: 60/126773
, PRIOR FILING DATE: 1999-03-29
, PRIOR APPLICATION NUMBER: 60/127887
, PRIOR FILING DATE: 1999-04-05
, PRIOR APPLICATION NUMBER: 60/130232
, PRIOR FILING DATE: 1999-04-21
, PRIOR APPLICATION NUMBER: 60/131022
, PRIOR FILING DATE: 1999-04-26
, PRIOR APPLICATION NUMBER: 60/131270
, PRIOR FILING DATE: 1999-04-27
, PRIOR APPLICATION NUMBER: 60/131291
, PRIOR FILING DATE: 1999-04-27
, PRIOR APPLICATION NUMBER: 60/131445
, PRIOR FILING DATE: 1999-04-28
, PRIOR APPLICATION NUMBER: 60/134287
, PRIOR FILING DATE: 1999-05-14
, PRIOR APPLICATION NUMBER: 60/140650
, PRIOR FILING DATE: 1999-06-22
, PRIOR APPLICATION NUMBER: 60/140723
, PRIOR FILING DATE: 1999-06-22
, PRIOR APPLICATION NUMBER: 60/141037
, PRIOR FILING DATE: 1999-06-23
, PRIOR APPLICATION NUMBER: 60/144758
, PRIOR FILING DATE: 1999-07-20
, PRIOR APPLICATION NUMBER: 60/145698
, PRIOR FILING DATE: 1999-07-26
, PRIOR APPLICATION NUMBER: 60/146222
, PRIOR FILING DATE: 1999-07-28
, PRIOR APPLICATION NUMBER: 60/146963
, PRIOR FILING DATE: 1999-08-03
, PRIOR APPLICATION NUMBER: 60/149320
, PRIOR FILING DATE: 1999-08-17
, PRIOR APPLICATION NUMBER: 60/149638
, PRIOR FILING DATE: 1999-08-17
, PRIOR APPLICATION NUMBER: 60/151733
, PRIOR FILING DATE: 1999-08-31
, PRIOR APPLICATION NUMBER: 60/164418
, PRIOR FILING DATE: 1999-11-09
, PRIOR APPLICATION NUMBER: 60/166361
, PRIOR FILING DATE: 1999-11-16
, PRIOR APPLICATION NUMBER: 60/169445
, PRIOR FILING DATE: 1999-12-07
, PRIOR APPLICATION NUMBER: 60/169495
, PRIOR FILING DATE: 1999-12-07
, PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA 3762
Db 1729 AATGAAAAA 1776

RESULT 48

US-10-219-076-39
; Sequence 39, Application US/10219076
; Publication No. US20030078379A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P162
; CURRENT APPLICATION NUMBER: US/10/219,076
; CURRENT FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/095302
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26

US-10-219-076-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-076-39

QY 3715 AATGAAAAA 3762
Db 1729 AATGAAAAA 1776

RESULT 49

US-10-230-434-39
; Sequence 39, Application US/10230434
; Publication No. US20030078380A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.

;; PRIOR APPLICATION NUMBER: 60/098544
;; PRIOR FILING DATE: 1998-08-31
;; PRIOR APPLICATION NUMBER: 60/099596
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099598
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099803
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099811
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100038
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101786
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101922
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/106178
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106248
;; PRIOR FILING DATE: 1998-10-29
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106905
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108787
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108801
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108849
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112422
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113605
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775

;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Fred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 50
US-10-219-003-39
; Sequence 39, Application US/10219003
; Publication No. US2003008063A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC

TITLE OF INVENTION: ACIDS ENCODING THE SAME
FILE REFERENCE: P3530P1C12
CURRENT APPLICATION NUMBER: US/10/219,003
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 10/113,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-06
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/095302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803

PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29
PRIOR APPLICATION NUMBER: 60/127887
PRIOR FILING DATE: 1999-04-05
PRIOR APPLICATION NUMBER: 60/130232
PRIOR FILING DATE: 1999-04-21

RESULT 51
US-10-219-075-39
Sequence 39, Application US/10219075
Publication No. US20030088064A1
GENERAL INFORMATION:
APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
ACIDS ENCODING THE SAME
FILE REFERENCE: P3530P1C61
CURRENT APPLICATION NUMBER: US/10/219,075
CURRENT FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113

GENERAL INFORMATION:

APPLICANT: Baker, Kevin P.
APPLICANT: Desnoyers, Luc
APPLICANT: Gerritsen, Mary
APPLICANT: Goddard, Audrey
APPLICANT: Godowski, Paul J.
APPLICANT: Grimaldi, J. Christopher
APPLICANT: Gurney, Austin L.
APPLICANT: Smith, Victoria
APPLICANT: Stephan, Jean-Philippe P.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William L.

TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530PIC57
CURRENT APPLICATION NUMBER: US/10/219,464
CURRENT FILING DATE: 2002-08-14
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26

;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-219-464-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762

Db 1729 AATGAA 1776

RESULT 53
US-10-219-466-39
; Sequence 39, Application US/10219466
; Publication No. US2003008068A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530P1C47
; CURRENT APPLICATION NUMBER: US/10/219,466
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-466-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762

Db 1729 AATGAA 1776

RESULT 54
US-10-219-479-39
; Sequence 39, Application US/10219479
; Publication No. US2003008067A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C33
; CURRENT APPLICATION NUMBER: US/10/219,479
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-479-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762

Db 1729 AATGAA 1776

RESULT 55
US-10-219-481-39
; Sequence 39, Application US/10219481
; Publication No. US2003008068A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher

```
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC28
; CURRENT APPLICATION NUMBER: US/10/219,481
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-481-39

Query Match          1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 56
US-10-230-260-39
; Sequence 39, Application US/10230260
; Publication No. US20030088070A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC28
; CURRENT APPLICATION NUMBER: US/10/230,260
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
```

```
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-260-39

Query Match          1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 57
US-10-232-231-39
; Sequence 39, Application US/10232231
; Publication No. US20030088071A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC104
; CURRENT APPLICATION NUMBER: US/10/232,231
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
```

; Remaining Prior Application data removed - See File Wrapper or PALM.

; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-232-231-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762
Db 1729 AATGAAA 1776

RESULT 58

US-10-232-233-39
; Sequence 39, Application US/10232233
; Publication No. US2003008072A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530P1C108

; CURRENT APPLICATION NUMBER: US/10/232,233
; CURRENT FILING DATE: 2002-08-29

; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-232-233-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762
Db 1729 AATGAAA 1776

RESULT 59

US-10-216-165-39
; Sequence 39, Application US/10216165
; Publication No. US20030092886A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530P1C7

; CURRENT APPLICATION NUMBER: US/10/216,165
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-216-165-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAA 3762
Db 1729 AATGAAA 1776

RESULT 60

US-10-218-956-39
; Sequence 39, Application US/10218956
; Publication No. US20030092887A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria

APPLICANT: Stephan, Jean-Philippe F.
APPLICANT: Watanabe, Colin L.
APPLICANT: Wood, William I.
TITLE OF INVENTION: ACIDIC AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
FILE REFERENCE: P3530P1C15
CURRENT APPLICATION NUMBER: US/10/218,956
CURRENT FILING DATE: 2002-08-12
PRIOR APPLICATION NUMBER: 10/119,480
PRIOR FILING DATE: 2002-04-09
PRIOR APPLICATION NUMBER: 60/059113
PRIOR FILING DATE: 1997-09-17
PRIOR APPLICATION NUMBER: 60/062287
PRIOR FILING DATE: 1997-10-17
PRIOR APPLICATION NUMBER: 60/063549
PRIOR FILING DATE: 1997-10-28
PRIOR APPLICATION NUMBER: 60/064103
PRIOR FILING DATE: 1997-10-31
PRIOR APPLICATION NUMBER: 60/069873
PRIOR FILING DATE: 1997-12-17
PRIOR APPLICATION NUMBER: 60/078910
PRIOR FILING DATE: 1998-03-20
PRIOR APPLICATION NUMBER: 60/079294
PRIOR FILING DATE: 1998-03-25
PRIOR APPLICATION NUMBER: 60/079656
PRIOR FILING DATE: 1998-03-26
PRIOR APPLICATION NUMBER: 60/079728
PRIOR FILING DATE: 1998-03-27
PRIOR APPLICATION NUMBER: 60/081819
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/081955
PRIOR FILING DATE: 1998-04-15
PRIOR APPLICATION NUMBER: 60/082804
PRIOR FILING DATE: 1998-04-22
PRIOR APPLICATION NUMBER: 60/084441
PRIOR FILING DATE: 1998-05-05
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/086392
PRIOR FILING DATE: 1998-05-22
PRIOR APPLICATION NUMBER: 60/089532
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089538
PRIOR FILING DATE: 1998-06-17
PRIOR APPLICATION NUMBER: 60/089905
PRIOR FILING DATE: 1998-06-18
PRIOR APPLICATION NUMBER: 60/090472
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090557
PRIOR FILING DATE: 1998-06-24
PRIOR APPLICATION NUMBER: 60/090691
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/090695
PRIOR FILING DATE: 1998-06-25
PRIOR APPLICATION NUMBER: 60/091982
PRIOR FILING DATE: 1998-07-07
PRIOR APPLICATION NUMBER: 60/093302
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/093318
PRIOR FILING DATE: 1998-08-04
PRIOR APPLICATION NUMBER: 60/095916
PRIOR FILING DATE: 1998-08-10
PRIOR APPLICATION NUMBER: 60/096146
PRIOR FILING DATE: 1998-08-11
PRIOR APPLICATION NUMBER: 60/096791
PRIOR FILING DATE: 1998-08-17
PRIOR APPLICATION NUMBER: 60/097986
PRIOR FILING DATE: 1998-08-26
PRIOR APPLICATION NUMBER: 60/098544
PRIOR FILING DATE: 1998-08-31
PRIOR APPLICATION NUMBER: 60/099596
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099598
PRIOR FILING DATE: 1998-09-09
PRIOR APPLICATION NUMBER: 60/099803
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099811
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099812
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/099816
PRIOR FILING DATE: 1998-09-10
PRIOR APPLICATION NUMBER: 60/100038
PRIOR FILING DATE: 1998-09-11
PRIOR APPLICATION NUMBER: 60/100385
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100390
PRIOR FILING DATE: 1998-09-15
PRIOR APPLICATION NUMBER: 60/100627
PRIOR FILING DATE: 1998-09-16
PRIOR APPLICATION NUMBER: 60/100848
PRIOR FILING DATE: 1998-09-18
PRIOR APPLICATION NUMBER: 60/100919
PRIOR FILING DATE: 1998-09-17
PRIOR APPLICATION NUMBER: 60/101477
PRIOR FILING DATE: 1998-09-23
PRIOR APPLICATION NUMBER: 60/101738
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101741
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101786
PRIOR FILING DATE: 1998-09-25
PRIOR APPLICATION NUMBER: 60/101916
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/101922
PRIOR FILING DATE: 1998-09-24
PRIOR APPLICATION NUMBER: 60/106178
PRIOR FILING DATE: 1998-10-28
PRIOR APPLICATION NUMBER: 60/106248
PRIOR FILING DATE: 1998-10-29
PRIOR APPLICATION NUMBER: 60/106464
PRIOR FILING DATE: 1998-10-30
PRIOR APPLICATION NUMBER: 60/106905
PRIOR FILING DATE: 1998-11-03
PRIOR APPLICATION NUMBER: 60/108787
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108801
PRIOR FILING DATE: 1998-11-17
PRIOR APPLICATION NUMBER: 60/108849
PRIOR FILING DATE: 1998-11-18
PRIOR APPLICATION NUMBER: 60/112422
PRIOR FILING DATE: 1998-12-15
PRIOR APPLICATION NUMBER: 60/113296
PRIOR FILING DATE: 1998-12-22
PRIOR APPLICATION NUMBER: 60/113605
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/113621
PRIOR FILING DATE: 1998-12-23
PRIOR APPLICATION NUMBER: 60/115558
PRIOR FILING DATE: 1998-01-12
PRIOR APPLICATION NUMBER: 60/115565
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/115733
PRIOR FILING DATE: 1999-01-12
PRIOR APPLICATION NUMBER: 60/119549
PRIOR FILING DATE: 1999-02-10
PRIOR APPLICATION NUMBER: 60/123618
PRIOR FILING DATE: 1999-03-10
PRIOR APPLICATION NUMBER: 60/125259
PRIOR FILING DATE: 1999-03-19
PRIOR APPLICATION NUMBER: 60/125775
PRIOR FILING DATE: 1999-03-23
PRIOR APPLICATION NUMBER: 60/126773
PRIOR FILING DATE: 1999-03-29

;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 61
US-10-219-468-39
; Sequence 39, Application US/10219468
; Publication No. US20030092888A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC34
; CURRENT APPLICATION NUMBER: US/10/219,468

;; CURRENT FILING DATE: 2002-08-13
;; PRIOR APPLICATION NUMBER: 10/119,480
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-219-468-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAA 3762
Db 1729 AATGAA 1776

RESULT 62
US-10-219-478-39
; Sequence 39, Application US/10219478
; Publication No. US20030092889A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC30
; CURRENT APPLICATION NUMBER: US/10/219,478
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20

```

; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-478-39

```

```
Query Match      1.3%; Score 48; DB 15; Length 1937;
Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 40; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```

RESULT 63
US-10-219-536-39
/ Sequence 39, Application US/10219536
/ Publication No. US2003092890A1
/ GENERAL INFORMATION:
/ APPLICANT: Baker, Kevin P.
/ APPLICANT: Desnoyers, Luc
/ APPLICANT: Gerritsen, Mary
/ APPLICANT: Goddard, Audrey
/ APPLICANT: Godowski, Paul J.
/ APPLICANT: Grimaldi, J. Christopher
/ APPLICANT: Gurney, Austin L.
/ APPLICANT: Smith, Victoria
/ APPLICANT: Stephan, Jean-Philippe F.
/ APPLICANT: Watanabe, Colin L.
/ APPLICANT: Wood, William I.
/ TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
/ TITLE OF INVENTION: ACIDS ENCODING THE SAME
/ FILE REFERENCE: P3530PIC67
/ CURRENT APPLICATION NUMBER: US/10/219,536

Query Match
1.3%; Score 48; DB 15; Length 1837;

```

Best Local Similarity 100.0%; Pred No. 3.2e-11; Gaps 0;
Matches 48; Conservative 0; Mismatches 0; Indels 0;

Qy 3715 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 1729 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
    |||||

RESULT 64
US-10-233-205-39
; Sequence 39, Application US/10233205
; Publication No. US20030096362A1
; GENERAL INFORMATION:

```

RESULT 64
US-10-233-205-39
; Sequence 39, Application US/10233205
; Publication No. US20030096362A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.

Query Match	1.3%	Score 48;	DB 15;	Length 1837;
Best Local Similarity	100.0%;	Pred. No. 3.2e-11;		
Matches 48;	Conservative	0;	Mismatches 0;	Indels 0;
Gaps	0;			

QY

3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
|||||

Dβ

1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 65
US-10-219-072-39
; Sequence 39, Application US/10219072
; Publication No. US20030096959A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Dennoyers, Luc

```

; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-470-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 67
US-10-219-474-39
; Sequence 39, Application US/10219474
; Publication No. US20030096961A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Deanovers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530PIC36
; CURRENT APPLICATION NUMBER: US/10/219,474
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25

```

```

; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-524-39
Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 68
US-10-219-524-39
; Sequence 39, Application US/10219524
; Publication No. US20030096962A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C37
; CURRENT APPLICATION NUMBER: US/10/219,524
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25

```

```

Qy 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 69
US-10-219-528-39
; Sequence 39, Application US/10219528
; Publication No. US20030096963A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C38
; CURRENT APPLICATION NUMBER: US/10/219,528
; PRIOR FILING DATE: 2002-08-14
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25

```

;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/095302
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095318
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095916
;; PRIOR FILING DATE: 1998-08-10
;; PRIOR APPLICATION NUMBER: 60/096146
;; PRIOR FILING DATE: 1998-08-11
;; PRIOR APPLICATION NUMBER: 60/096791
;; PRIOR FILING DATE: 1998-08-17
;; PRIOR APPLICATION NUMBER: 60/097986
;; PRIOR FILING DATE: 1998-08-26
;; PRIOR APPLICATION NUMBER: 60/098544
;; PRIOR FILING DATE: 1998-08-31
;; PRIOR APPLICATION NUMBER: 60/099596
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099598
;; PRIOR FILING DATE: 1998-09-09
;; PRIOR APPLICATION NUMBER: 60/099803
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099811
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099812
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/099816
;; PRIOR FILING DATE: 1998-09-10
;; PRIOR APPLICATION NUMBER: 60/100038
;; PRIOR FILING DATE: 1998-09-11
;; PRIOR APPLICATION NUMBER: 60/100385
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100390
;; PRIOR FILING DATE: 1998-09-15
;; PRIOR APPLICATION NUMBER: 60/100627
;; PRIOR FILING DATE: 1998-09-16
;; PRIOR APPLICATION NUMBER: 60/100848
;; PRIOR FILING DATE: 1998-09-18
;; PRIOR APPLICATION NUMBER: 60/100919
;; PRIOR FILING DATE: 1998-09-17
;; PRIOR APPLICATION NUMBER: 60/101477
;; PRIOR FILING DATE: 1998-09-23
;; PRIOR APPLICATION NUMBER: 60/101738
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101741
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101786
;; PRIOR FILING DATE: 1998-09-25
;; PRIOR APPLICATION NUMBER: 60/101916
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/101922
;; PRIOR FILING DATE: 1998-09-24
;; PRIOR APPLICATION NUMBER: 60/106178
;; PRIOR FILING DATE: 1998-10-28
;; PRIOR APPLICATION NUMBER: 60/106248
;; PRIOR FILING DATE: 1998-10-29
;; PRIOR APPLICATION NUMBER: 60/106464
;; PRIOR FILING DATE: 1998-10-30
;; PRIOR APPLICATION NUMBER: 60/106905
;; PRIOR FILING DATE: 1998-11-03
;; PRIOR APPLICATION NUMBER: 60/108787
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108801
;; PRIOR FILING DATE: 1998-11-17
;; PRIOR APPLICATION NUMBER: 60/108849
;; PRIOR FILING DATE: 1998-11-18
;; PRIOR APPLICATION NUMBER: 60/112422
;; PRIOR FILING DATE: 1998-12-15
;; PRIOR APPLICATION NUMBER: 60/113296
;; PRIOR FILING DATE: 1998-12-22
;; PRIOR APPLICATION NUMBER: 60/113605

;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/113621
;; PRIOR FILING DATE: 1998-12-23
;; PRIOR APPLICATION NUMBER: 60/115558
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115565
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 70

US-10-227-880-39
; Sequence 39, Application US/10227880
; Publication No. US20030096964A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gunney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: ACIDS AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C74
; CURRENT APPLICATION NUMBER: US/10/227,880
; CURRENT FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/093302
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12

;; PRIOR APPLICATION NUMBER: 60/115733
;; PRIOR FILING DATE: 1999-01-12
;; PRIOR APPLICATION NUMBER: 60/119549
;; PRIOR FILING DATE: 1999-02-10
;; PRIOR APPLICATION NUMBER: 60/123618
;; PRIOR FILING DATE: 1999-03-10
;; PRIOR APPLICATION NUMBER: 60/125259
;; PRIOR FILING DATE: 1999-03-19
;; PRIOR APPLICATION NUMBER: 60/125775
;; PRIOR FILING DATE: 1999-03-23
;; PRIOR APPLICATION NUMBER: 60/126773
;; PRIOR FILING DATE: 1999-03-29
;; PRIOR APPLICATION NUMBER: 60/127887
;; PRIOR FILING DATE: 1999-04-05
;; PRIOR APPLICATION NUMBER: 60/130232
;; PRIOR FILING DATE: 1999-04-21
;; PRIOR APPLICATION NUMBER: 60/131022
;; PRIOR FILING DATE: 1999-04-26
;; PRIOR APPLICATION NUMBER: 60/131270
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131291
;; PRIOR FILING DATE: 1999-04-27
;; PRIOR APPLICATION NUMBER: 60/131445
;; PRIOR FILING DATE: 1999-04-28
;; PRIOR APPLICATION NUMBER: 60/134287
;; PRIOR FILING DATE: 1999-05-14
;; PRIOR APPLICATION NUMBER: 60/140650
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/140723
;; PRIOR FILING DATE: 1999-06-22
;; PRIOR APPLICATION NUMBER: 60/141037
;; PRIOR FILING DATE: 1999-06-23
;; PRIOR APPLICATION NUMBER: 60/144758
;; PRIOR FILING DATE: 1999-07-20
;; PRIOR APPLICATION NUMBER: 60/145698
;; PRIOR FILING DATE: 1999-07-26
;; PRIOR APPLICATION NUMBER: 60/146222
;; PRIOR FILING DATE: 1999-07-28
;; PRIOR APPLICATION NUMBER: 60/146963
;; PRIOR FILING DATE: 1999-08-03
;; PRIOR APPLICATION NUMBER: 60/149320
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/149638
;; PRIOR FILING DATE: 1999-08-17
;; PRIOR APPLICATION NUMBER: 60/151733
;; PRIOR FILING DATE: 1999-08-31
;; PRIOR APPLICATION NUMBER: 60/164418
;; PRIOR FILING DATE: 1999-11-09
;; PRIOR APPLICATION NUMBER: 60/166361
;; PRIOR FILING DATE: 1999-11-16
;; PRIOR APPLICATION NUMBER: 60/169445
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169495
;; PRIOR FILING DATE: 1999-12-07
;; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred.No. 3.2e-11;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
DB 1729 AATGAA 1776

RESULT 71

US-10-227-881-39

; Sequence 39, Application US/10227881

; Publication No. US20030096965A1

; GENERAL INFORMATION:

; APPLICANT: Baker, Kevin P.

; APPLICANT: Desnoyers, Luc

; APPLICANT: Gerritsen, Mary

;; APPLICANT: Goddard, Audrey
;; APPLICANT: Godowski, Paul J.
;; APPLICANT: Grimaldi, J. Christopher
;; APPLICANT: Gurney, Austin L.
;; APPLICANT: Smith, Victoria
;; APPLICANT: Stephan, Jean-Philippe F.
;; APPLICANT: Watanabe, Colin L.
;; APPLICANT: Wood, William I.
;; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
;; FILE REFERENCE: P3530PLC80
;; CURRENT APPLICATION NUMBER: US/10/227,881
;; PRIOR FILING DATE: 2002-08-26
;; PRIOR APPLICATION NUMBER: 10/119,480
;; PRIOR FILING DATE: 2002-04-09
;; PRIOR APPLICATION NUMBER: 60/059113
;; PRIOR FILING DATE: 1997-09-17
;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; PRIOR APPLICATION NUMBER: 60/081819
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/081955
;; PRIOR FILING DATE: 1998-04-15
;; PRIOR APPLICATION NUMBER: 60/082804
;; PRIOR FILING DATE: 1998-04-22
;; PRIOR APPLICATION NUMBER: 60/084441
;; PRIOR FILING DATE: 1998-05-06
;; PRIOR APPLICATION NUMBER: 60/085323
;; PRIOR FILING DATE: 1998-05-13
;; PRIOR APPLICATION NUMBER: 60/085579
;; PRIOR FILING DATE: 1998-05-15
;; PRIOR APPLICATION NUMBER: 60/086392
;; PRIOR FILING DATE: 1998-05-22
;; PRIOR APPLICATION NUMBER: 60/089532
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089538
;; PRIOR FILING DATE: 1998-06-17
;; PRIOR APPLICATION NUMBER: 60/089905
;; PRIOR FILING DATE: 1998-06-18
;; PRIOR APPLICATION NUMBER: 60/090472
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090557
;; PRIOR FILING DATE: 1998-06-24
;; PRIOR APPLICATION NUMBER: 60/090691
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/090695
;; PRIOR FILING DATE: 1998-06-25
;; PRIOR APPLICATION NUMBER: 60/091982
;; PRIOR FILING DATE: 1998-07-07
;; PRIOR APPLICATION NUMBER: 60/095302
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095318
;; PRIOR FILING DATE: 1998-08-04
;; PRIOR APPLICATION NUMBER: 60/095916
;; PRIOR FILING DATE: 1998-08-10
;; PRIOR APPLICATION NUMBER: 60/096146
;; PRIOR FILING DATE: 1998-08-11
;; PRIOR APPLICATION NUMBER: 60/096791
;; PRIOR FILING DATE: 1998-08-17

; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-16
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259

; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887
; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
Db 1729 AATGAA 1776
|||||

RESULT 72

US-10-227-882-39
; Sequence 39, Application US/10227882
; Publication No. US20030096966A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.


```
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-15
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835
```

```
Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 3715 ATGAAAAA 3762
Db 1729 ATGAAAAA 1776
```

```
RESULT 73
US-10-230-436-39
; Sequence 39, Application US/10230436
; Publication No. US20030096761
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C97
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
```

```
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-436-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11; Mismatches 0; Indels 0; Gaps 0;
Matches 48; Conservative 0;

QY 3715 ATGAAAAA 3762
Db 1729 ATGAAAAA 1776
```

```
RESULT 74
US-10-232-223-39
; Sequence 39, Application US/10232223
; Publication No. US20030096681
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C102
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
```

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels

```
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC105
; CURRENT APPLICATION NUMBER: US/10/232,229
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC106
; CURRENT APPLICATION NUMBER: US/10/232,234
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC106
; CURRENT APPLICATION NUMBER: US/10/232,234
; CURRENT FILING DATE: 2002-08-29
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
```

```
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
; US-10-232-234-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 79
US-10-219-060-39
; Sequence 39, Application US/10219060
; Publication No. US20030100064A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC22
; CURRENT APPLICATION NUMBER: US/10/219,060
; CURRENT FILING DATE: 2002-08-12
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
```

1	PRIOR FILING DATE: 1998-09-24
2	PRIOR APPLICATION NUMBER: 60/101741
3	PRIOR FILING DATE: 1998-09-24
4	PRIOR APPLICATION NUMBER: 60/101786
5	PRIOR FILING DATE: 1998-09-25
6	PRIOR APPLICATION NUMBER: 60/101916
7	PRIOR FILING DATE: 1998-09-24
8	PRIOR APPLICATION NUMBER: 60/101922
9	PRIOR FILING DATE: 1998-09-24
10	PRIOR APPLICATION NUMBER: 60/106178
11	PRIOR FILING DATE: 1998-10-28
12	PRIOR APPLICATION NUMBER: 60/106248
13	PRIOR FILING DATE: 1998-10-29
14	PRIOR APPLICATION NUMBER: 60/106464
15	PRIOR FILING DATE: 1998-10-30
16	PRIOR APPLICATION NUMBER: 60/106905
17	PRIOR FILING DATE: 1998-11-03
18	PRIOR APPLICATION NUMBER: 60/108787
19	PRIOR FILING DATE: 1998-11-17
20	PRIOR APPLICATION NUMBER: 60/108801
21	PRIOR FILING DATE: 1998-11-17
22	PRIOR APPLICATION NUMBER: 60/108949
23	PRIOR FILING DATE: 1998-11-18
24	PRIOR APPLICATION NUMBER: 60/112422
25	PRIOR FILING DATE: 1998-12-15
26	PRIOR APPLICATION NUMBER: 60/113296
27	PRIOR FILING DATE: 1998-12-22
28	PRIOR APPLICATION NUMBER: 60/113605
29	PRIOR FILING DATE: 1998-12-23
30	PRIOR APPLICATION NUMBER: 60/113621
31	PRIOR FILING DATE: 1998-12-23
32	PRIOR APPLICATION NUMBER: 60/115558
33	PRIOR FILING DATE: 1999-01-12
34	PRIOR APPLICATION NUMBER: 60/115565
35	PRIOR FILING DATE: 1999-01-12
36	PRIOR APPLICATION NUMBER: 60/115733
37	PRIOR FILING DATE: 1999-01-12
38	PRIOR APPLICATION NUMBER: 60/119549
39	PRIOR FILING DATE: 1999-02-10
40	PRIOR APPLICATION NUMBER: 60/123618
41	PRIOR FILING DATE: 1999-03-10
42	PRIOR APPLICATION NUMBER: 60/125259
43	PRIOR FILING DATE: 1999-03-19
44	PRIOR APPLICATION NUMBER: 60/125775
45	PRIOR FILING DATE: 1999-03-23
46	PRIOR APPLICATION NUMBER: 60/126773
47	PRIOR FILING DATE: 1999-03-29
48	PRIOR APPLICATION NUMBER: 60/127887
49	PRIOR FILING DATE: 1999-04-05
50	PRIOR APPLICATION NUMBER: 60/130232
51	PRIOR FILING DATE: 1999-04-21
52	PRIOR APPLICATION NUMBER: 60/131022
53	PRIOR FILING DATE: 1999-04-26
54	PRIOR APPLICATION NUMBER: 60/131270
55	PRIOR FILING DATE: 1999-04-27
56	PRIOR APPLICATION NUMBER: 60/131291
57	PRIOR FILING DATE: 1999-04-27
58	PRIOR APPLICATION NUMBER: 60/131445
59	PRIOR FILING DATE: 1999-04-28
60	PRIOR APPLICATION NUMBER: 60/134287
61	PRIOR FILING DATE: 1999-05-14
62	PRIOR APPLICATION NUMBER: 60/140650
63	PRIOR FILING DATE: 1999-06-22
64	PRIOR APPLICATION NUMBER: 60/140723
65	PRIOR FILING DATE: 1999-06-22
66	PRIOR APPLICATION NUMBER: 60/141037
67	PRIOR FILING DATE: 1999-06-23
68	PRIOR APPLICATION NUMBER: 60/144758
69	PRIOR FILING DATE: 1999-07-20
70	PRIOR APPLICATION NUMBER: 60/145698
71	PRIOR FILING DATE: 1999-07-26
72	PRIOR APPLICATION NUMBER: 60/146222
73	PRIOR FILING DATE: 1999-07-28

; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
DB 1729 AATGAA 1776

RESULT 80
US-10-216-160-39
; Sequence 39, Application US/10216160
; Publication No. US20030100708A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC8
; CURRENT APPLICATION NUMBER: US/10/216.160
; CURRENT FILING DATE: 2002-08-09
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-216-160-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
DB 1729 AATGAA 1776

RESULT 81
US-10-216-162-39
; Sequence 39, Application US/10216162
; Publication No. US20030100709A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary

; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC2
; CURRENT APPLICATION NUMBER: US/10/216.162
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/119.480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-216-162-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
|||||
DB 1729 AATGAA 1776

RESULT 82
US-10-216-164-39
; Sequence 39, Application US/10216164
; Publication No. US20030100710A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC5
; CURRENT APPLICATION NUMBER: US/10/216.164
; CURRENT FILING DATE: 2002-08-09
; PRIOR APPLICATION NUMBER: 10/119.480
; PRIOR FILING DATE: 2002-04-09

```
Query Match      1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```


;; PRIOR APPLICATION NUMBER: 60/062287
;; PRIOR FILING DATE: 1997-10-17
;; PRIOR APPLICATION NUMBER: 60/063549
;; PRIOR FILING DATE: 1997-10-28
;; PRIOR APPLICATION NUMBER: 60/064103
;; PRIOR FILING DATE: 1997-10-31
;; PRIOR APPLICATION NUMBER: 60/069873
;; PRIOR FILING DATE: 1997-12-17
;; PRIOR APPLICATION NUMBER: 60/078910
;; PRIOR FILING DATE: 1998-03-20
;; PRIOR APPLICATION NUMBER: 60/079294
;; PRIOR FILING DATE: 1998-03-25
;; PRIOR APPLICATION NUMBER: 60/079656
;; PRIOR FILING DATE: 1998-03-26
;; PRIOR APPLICATION NUMBER: 60/079728
;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-219-074-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;
QY 3715 AATGAAAAA 3762
Db 1729 AATGAAAAA 1776

RESULT 88
US-10-219-077-39
; Sequence 39, Application US/10219077
; Publication No. US20030100716A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Destoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe P.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC53
; CURRENT APPLICATION NUMBER: US/10/219,077
; PRIOR APPLICATION NUMBER: 2002-08-13
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-077-39

;; PRIOR FILING DATE: 1998-03-27
;; Remaining Prior Application data removed - See File Wrapper or PALM.
;; NUMBER OF SEQ ID NOS: 246
;; SEQ ID NO 39
;; LENGTH: 1837
;; TYPE: DNA
;; ORGANISM: Homo Sapien
US-10-219-077-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;
QY 3715 AATGAAAAA 3762
Db 1729 AATGAAAAA 1776

RESULT 89
US-10-219-465-39
; Sequence 39, Application US/10219465
; Publication No. US20030100717A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Destoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe P.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC50
; CURRENT APPLICATION NUMBER: US/10/219,465
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-465-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11; Indels 0; Gaps 0;
Matches 48; Conservative 0; Mismatches 0;
QY 3715 AATGAAAAA 3762


```
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-471-39

Query Match
Best Local Similarity 1.3%; Score 48; DB 15; Length 1837;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 93
US-10-219-473-39
; Sequence 39, Application US/10219473
; Publication NO. US20030100721A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C35
; CURRENT APPLICATION NUMBER: US/10/219,473
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-471-39

Query Match
Best Local Similarity 1.3%; Score 48; DB 15; Length 1837;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 93
US-10-219-473-39
; Sequence 39, Application US/10219473
; Publication NO. US20030100721A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C35
; CURRENT APPLICATION NUMBER: US/10/219,473
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
```

```
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-473-39

Query Match
Best Local Similarity 1.3%; Score 48; DB 15; Length 1837;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 94
US-10-219-476-39
; Sequence 39, Application US/10219476
; Publication NO. US20030100722A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530P1C26
; CURRENT APPLICATION NUMBER: US/10/219,476
; CURRENT FILING DATE: 2002-08-13
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-219-476-39

Query Match
Best Local Similarity 1.3%; Score 48; DB 15; Length 1837;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
```


; PRIOR APPLICATION NUMBER: 60/093958
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/093803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/093811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/093812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/093816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100390
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100627
; PRIOR FILING DATE: 1998-09-15
; PRIOR APPLICATION NUMBER: 60/100848
; PRIOR FILING DATE: 1998-09-18
; PRIOR APPLICATION NUMBER: 60/100919
; PRIOR FILING DATE: 1998-09-17
; PRIOR APPLICATION NUMBER: 60/101477
; PRIOR FILING DATE: 1998-09-23
; PRIOR APPLICATION NUMBER: 60/101738
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101741
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101786
; PRIOR FILING DATE: 1998-09-25
; PRIOR APPLICATION NUMBER: 60/101916
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/101922
; PRIOR FILING DATE: 1998-09-24
; PRIOR APPLICATION NUMBER: 60/106178
; PRIOR FILING DATE: 1998-10-28
; PRIOR APPLICATION NUMBER: 60/106248
; PRIOR FILING DATE: 1998-10-29
; PRIOR APPLICATION NUMBER: 60/106464
; PRIOR FILING DATE: 1998-10-30
; PRIOR APPLICATION NUMBER: 60/106905
; PRIOR FILING DATE: 1998-11-03
; PRIOR APPLICATION NUMBER: 60/108787
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108801
; PRIOR FILING DATE: 1998-11-17
; PRIOR APPLICATION NUMBER: 60/108849
; PRIOR FILING DATE: 1998-11-18
; PRIOR APPLICATION NUMBER: 60/112422
; PRIOR FILING DATE: 1998-12-15
; PRIOR APPLICATION NUMBER: 60/113296
; PRIOR FILING DATE: 1998-12-22
; PRIOR APPLICATION NUMBER: 60/113605
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/113621
; PRIOR FILING DATE: 1998-12-23
; PRIOR APPLICATION NUMBER: 60/115558
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115565
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/115733
; PRIOR FILING DATE: 1999-01-12
; PRIOR APPLICATION NUMBER: 60/119549
; PRIOR FILING DATE: 1999-02-10
; PRIOR APPLICATION NUMBER: 60/123618
; PRIOR FILING DATE: 1999-03-10
; PRIOR APPLICATION NUMBER: 60/125259
; PRIOR FILING DATE: 1999-03-19
; PRIOR APPLICATION NUMBER: 60/125775
; PRIOR FILING DATE: 1999-03-23
; PRIOR APPLICATION NUMBER: 60/126773
; PRIOR FILING DATE: 1999-03-29
; PRIOR APPLICATION NUMBER: 60/127887

; PRIOR FILING DATE: 1999-04-05
; PRIOR APPLICATION NUMBER: 60/130232
; PRIOR FILING DATE: 1999-04-21
; PRIOR APPLICATION NUMBER: 60/131022
; PRIOR FILING DATE: 1999-04-26
; PRIOR APPLICATION NUMBER: 60/131270
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
DB 1729 AATGAA 1776

RESULT 97

US-10-227-876-39
; Sequence 39, Application US/10227876
; Publication No. US20030100725A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE REFERENCE: P3530PIC73
; CURRENT APPLICATION NUMBER: US/10/227,876
; CURRENT FILING DATE: 2002-08-26

, PRIOR APPLICATION NUMBER: 10/119,480
, PRIOR FILING DATE: 2002-04-09
, PRIOR APPLICATION NUMBER: 60/059113
, PRIOR FILING DATE: 1997-09-17
, PRIOR APPLICATION NUMBER: 60/062287
, PRIOR FILING DATE: 1997-10-17
, PRIOR APPLICATION NUMBER: 60/063549
, PRIOR FILING DATE: 1997-10-28
, PRIOR APPLICATION NUMBER: 60/064103
, PRIOR FILING DATE: 1997-10-31
, PRIOR APPLICATION NUMBER: 60/069873
, PRIOR FILING DATE: 1997-12-17
, PRIOR APPLICATION NUMBER: 60/078910
, PRIOR FILING DATE: 1998-03-20
, PRIOR APPLICATION NUMBER: 60/079294
, PRIOR FILING DATE: 1998-03-25
, PRIOR APPLICATION NUMBER: 60/079656
, PRIOR FILING DATE: 1998-03-26
, PRIOR APPLICATION NUMBER: 60/079728
, PRIOR FILING DATE: 1998-03-27
, PRIOR APPLICATION NUMBER: 60/081819
, PRIOR FILING DATE: 1998-04-15
, PRIOR APPLICATION NUMBER: 60/081955
, PRIOR FILING DATE: 1998-04-15
, PRIOR APPLICATION NUMBER: 60/082804
, PRIOR FILING DATE: 1998-04-22
, PRIOR APPLICATION NUMBER: 60/084441
, PRIOR FILING DATE: 1998-05-06
, PRIOR APPLICATION NUMBER: 60/085323
, PRIOR FILING DATE: 1998-05-13
, PRIOR APPLICATION NUMBER: 60/085579
, PRIOR FILING DATE: 1998-05-15
, PRIOR APPLICATION NUMBER: 60/086392
, PRIOR FILING DATE: 1998-05-22
, PRIOR APPLICATION NUMBER: 60/089532
, PRIOR FILING DATE: 1998-06-17
, PRIOR APPLICATION NUMBER: 60/089538
, PRIOR FILING DATE: 1998-06-17
, PRIOR APPLICATION NUMBER: 60/089905
, PRIOR FILING DATE: 1998-06-18
, PRIOR APPLICATION NUMBER: 60/090472
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090557
, PRIOR FILING DATE: 1998-06-24
, PRIOR APPLICATION NUMBER: 60/090691
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/090695
, PRIOR FILING DATE: 1998-06-25
, PRIOR APPLICATION NUMBER: 60/091982
, PRIOR FILING DATE: 1998-07-07
, PRIOR APPLICATION NUMBER: 60/093302
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095318
, PRIOR FILING DATE: 1998-08-04
, PRIOR APPLICATION NUMBER: 60/095916
, PRIOR FILING DATE: 1998-08-10
, PRIOR APPLICATION NUMBER: 60/096146
, PRIOR FILING DATE: 1998-08-11
, PRIOR APPLICATION NUMBER: 60/096791
, PRIOR FILING DATE: 1998-08-17
, PRIOR APPLICATION NUMBER: 60/097986
, PRIOR FILING DATE: 1998-08-25
, PRIOR APPLICATION NUMBER: 60/098544
, PRIOR FILING DATE: 1998-08-31
, PRIOR APPLICATION NUMBER: 60/099596
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099598
, PRIOR FILING DATE: 1998-09-09
, PRIOR APPLICATION NUMBER: 60/099803
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099811
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099812
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/099816
, PRIOR FILING DATE: 1998-09-10
, PRIOR APPLICATION NUMBER: 60/100038
, PRIOR FILING DATE: 1998-09-11
, PRIOR APPLICATION NUMBER: 60/100385
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100390
, PRIOR FILING DATE: 1998-09-15
, PRIOR APPLICATION NUMBER: 60/100627
, PRIOR FILING DATE: 1998-09-16
, PRIOR APPLICATION NUMBER: 60/100848
, PRIOR FILING DATE: 1998-09-16
, PRIOR APPLICATION NUMBER: 60/100919
, PRIOR FILING DATE: 1998-09-17
, PRIOR APPLICATION NUMBER: 60/101477
, PRIOR FILING DATE: 1998-09-23
, PRIOR APPLICATION NUMBER: 60/101738
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101741
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101786
, PRIOR FILING DATE: 1998-09-25
, PRIOR APPLICATION NUMBER: 60/101916
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/101922
, PRIOR FILING DATE: 1998-09-24
, PRIOR APPLICATION NUMBER: 60/106178
, PRIOR FILING DATE: 1998-10-28
, PRIOR APPLICATION NUMBER: 60/106248
, PRIOR FILING DATE: 1998-10-29
, PRIOR APPLICATION NUMBER: 60/106464
, PRIOR FILING DATE: 1998-10-30
, PRIOR APPLICATION NUMBER: 60/106905
, PRIOR FILING DATE: 1998-11-03
, PRIOR APPLICATION NUMBER: 60/108787
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108801
, PRIOR FILING DATE: 1998-11-17
, PRIOR APPLICATION NUMBER: 60/108849
, PRIOR FILING DATE: 1998-11-18
, PRIOR APPLICATION NUMBER: 60/112422
, PRIOR FILING DATE: 1998-12-15
, PRIOR APPLICATION NUMBER: 60/113296
, PRIOR FILING DATE: 1998-12-22
, PRIOR APPLICATION NUMBER: 60/113605
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/113621
, PRIOR FILING DATE: 1998-12-23
, PRIOR APPLICATION NUMBER: 60/115558
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115565
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/115733
, PRIOR FILING DATE: 1999-01-12
, PRIOR APPLICATION NUMBER: 60/119549
, PRIOR FILING DATE: 1999-02-10
, PRIOR APPLICATION NUMBER: 60/123618
, PRIOR FILING DATE: 1999-03-10
, PRIOR APPLICATION NUMBER: 60/125259
, PRIOR FILING DATE: 1999-03-19
, PRIOR APPLICATION NUMBER: 60/125775
, PRIOR FILING DATE: 1999-03-23
, PRIOR APPLICATION NUMBER: 60/126773
, PRIOR FILING DATE: 1999-03-29
, PRIOR APPLICATION NUMBER: 60/127887
, PRIOR FILING DATE: 1999-04-05
, PRIOR APPLICATION NUMBER: 60/130232
, PRIOR FILING DATE: 1999-04-21
, PRIOR APPLICATION NUMBER: 60/131022
, PRIOR FILING DATE: 1999-04-26
, PRIOR APPLICATION NUMBER: 60/131270
, PRIOR FILING DATE: 1999-04-27

```

; PRIOR APPLICATION NUMBER: 60/131291
; PRIOR FILING DATE: 1999-04-27
; PRIOR APPLICATION NUMBER: 60/131445
; PRIOR FILING DATE: 1999-04-28
; PRIOR APPLICATION NUMBER: 60/134287
; PRIOR FILING DATE: 1999-05-14
; PRIOR APPLICATION NUMBER: 60/140650
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/140723
; PRIOR FILING DATE: 1999-06-22
; PRIOR APPLICATION NUMBER: 60/141037
; PRIOR FILING DATE: 1999-06-23
; PRIOR APPLICATION NUMBER: 60/144758
; PRIOR FILING DATE: 1999-07-20
; PRIOR APPLICATION NUMBER: 60/145698
; PRIOR FILING DATE: 1999-07-26
; PRIOR APPLICATION NUMBER: 60/146222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: 60/146963
; PRIOR FILING DATE: 1999-08-03
; PRIOR APPLICATION NUMBER: 60/149320
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/149638
; PRIOR FILING DATE: 1999-08-17
; PRIOR APPLICATION NUMBER: 60/151733
; PRIOR FILING DATE: 1999-08-31
; PRIOR APPLICATION NUMBER: 60/164418
; PRIOR FILING DATE: 1999-11-09
; PRIOR APPLICATION NUMBER: 60/166361
; PRIOR FILING DATE: 1999-11-16
; PRIOR APPLICATION NUMBER: 60/169445
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169495
; PRIOR FILING DATE: 1999-12-07
; PRIOR APPLICATION NUMBER: 60/169835

Query Match      1.38; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3715 AATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 1729 AATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 98
US-10-227-878-39
; Sequence 39, Application US/10237878
; Publication No. US20030100726A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Deenoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; FILE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: P3530F1C76
; CURRENT APPLICATION NUMBER: US/10/227,878
; PRIOR FILING DATE: 2002-08-26
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/086392
; PRIOR FILING DATE: 1998-05-22
; PRIOR APPLICATION NUMBER: 60/089532
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089538
; PRIOR FILING DATE: 1998-06-17
; PRIOR APPLICATION NUMBER: 60/089905
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: 60/090472
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090557
; PRIOR FILING DATE: 1998-06-24
; PRIOR APPLICATION NUMBER: 60/090691
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/090695
; PRIOR FILING DATE: 1998-06-25
; PRIOR APPLICATION NUMBER: 60/091982
; PRIOR FILING DATE: 1998-07-07
; PRIOR APPLICATION NUMBER: 60/095302
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095318
; PRIOR FILING DATE: 1998-08-04
; PRIOR APPLICATION NUMBER: 60/095916
; PRIOR FILING DATE: 1998-08-10
; PRIOR APPLICATION NUMBER: 60/096146
; PRIOR FILING DATE: 1998-08-11
; PRIOR APPLICATION NUMBER: 60/096791
; PRIOR FILING DATE: 1998-08-17
; PRIOR APPLICATION NUMBER: 60/097986
; PRIOR FILING DATE: 1998-08-26
; PRIOR APPLICATION NUMBER: 60/098544
; PRIOR FILING DATE: 1998-08-31
; PRIOR APPLICATION NUMBER: 60/099596
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099598
; PRIOR FILING DATE: 1998-09-09
; PRIOR APPLICATION NUMBER: 60/099803
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099811
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099812
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/099816
; PRIOR FILING DATE: 1998-09-10
; PRIOR APPLICATION NUMBER: 60/100038
; PRIOR FILING DATE: 1998-09-11
; PRIOR APPLICATION NUMBER: 60/100385
; PRIOR FILING DATE: 1998-09-15
```


; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-229-974-39

Query Match 1.3%; Score 48; DB 15; Length 1837;
Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3715 AATGAAA 3762
DB 1729 AATGAAA 1776

RESULT 100

US-10-230-024-39
; Sequence 39, Application US/10230024
; Publication No. US20030100728A1
; GENERAL INFORMATION:
; APPLICANT: Baker, Kevin P.
; APPLICANT: Desnoyers, Luc
; APPLICANT: Gerritsen, Mary
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Smith, Victoria
; APPLICANT: Stephan, Jean-Philippe F.
; APPLICANT: Watanabe, Colin L.
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: SECRETED AND TRANSMEMBRANE POLYPEPTIDES AND NUCLEIC
; TITLE OF INVENTION: ACIDS ENCODING THE SAME
; FILE REFERENCE: F3530P1C89
; CURRENT APPLICATION NUMBER: US/10/230.024
; CURRENT FILING DATE: 2002-08-28
; PRIOR APPLICATION NUMBER: 10/119,480
; PRIOR FILING DATE: 2002-04-09
; PRIOR APPLICATION NUMBER: 60/059113
; PRIOR FILING DATE: 1997-09-17
; PRIOR APPLICATION NUMBER: 60/062287
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/063549
; PRIOR FILING DATE: 1997-10-28
; PRIOR APPLICATION NUMBER: 60/064103
; PRIOR FILING DATE: 1997-10-31
; PRIOR APPLICATION NUMBER: 60/069873
; PRIOR FILING DATE: 1997-12-17
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 246
; SEQ ID NO 39
; LENGTH: 1837
; TYPE: DNA
; ORGANISM: Homo Sapien
US-10-230-024-39

Query Match

1.3%; Score 48; DB 15; Length 1837;

Best Local Similarity 100.0%; Pred.No. 3.2e-11;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
OY 3715 AATGAAA 3762
DB 1729 AATGAAA 1776

Search completed: April 23, 2004, 14:38:17
Job time : 1475 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 03:08:45 ; Search time 1339 seconds
(without alignments)

11935.565 Million cell updates/sec

Title: US-10-005-907-1

Perfect score: 3762

Sequence: 1 gagaaacggagtcactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC

Gapop_60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database : N_Geneseq_29Jan04.*
1: Geneseq1980s.*
2: Geneseq1990s.*
3: Geneseq2000s.*
4: Geneseq2001as.*
5: Geneseq2001bs.*
6: Geneseq2002s.*
7: Geneseq2003as.*
8: Geneseq2003bs.*
9: Geneseq2003cs.*
10: Geneseq2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3762	100.0	3762	6	ABN81319 Human mas
2	733	19.5	1128	5	ABV20298 Human pro
3	733	19.5	1128	5	ABV26130 Human pro
4	733	19.5	1128	5	ABV26458 Human pro
5	733	19.5	1128	5	ABV20617 Human pro
6	733	19.5	1128	5	ABV20806 Human pro
7	733	19.5	1128	5	ABV20439 Human pro
8	733	19.5	1128	5	ABV26275 Human pro
9	733	19.5	1128	5	ABV26523 Human pro
10	506	13.5	667	9	ADC30760 Human nov
11	344	9.1	397	5	ABV10439 Human pro
12	344	9.1	433	5	ABV01270 Human pro
13	342	9.1	437	5	ABV31609 Human pro
14	342	9.1	437	5	ABV40245 Human pro
15	342	9.1	437	5	ABV40577 Human pro
16	244	6.5	446	5	ABV00935 Human pro
17	213	5.7	404	5	ABV10104 Human pro
18	213	5.7	421	5	ABV31276 Human pro
19	204	5.4	359	5	ABV09593 Human pro
20	203	5.4	400	5	ABV00424 Human pro
21	203	5.4	401	5	ABV39737 Human pro
22	203	5.4	401	5	ABV30769 Human pro
23	203	5.4	401	5	ABV39958 Human pro

24	62	1.6	296	5	ABV30990 Human pro
25	62	1.6	337	5	ABV00647 Human pro
26	62	1.6	425	5	ABV09816 Human pro
27	52	1.4	381	5	ABV56428 Human pro
28	50	1.3	345	4	AAI90783 Human pol
29	50	1.3	424	5	ABV57393 Human pro
30	50	1.3	449	4	AAI83050 Human pol
31	50	1.3	1702	3	AAF15960 Human pro
32	50	1.3	1820	9	ADD18824 Human dis
33	50	1.3	3602	7	ACF12861 Human cer
34	50	1.3	3607	2	AAH87412 Hepatocel
35	49	1.3	265	5	ABV48235 Human pro
36	49	1.3	287	5	AAV18451 Human pro
37	49	1.3	381	4	AAI84445 Human pol
38	49	1.3	476	6	ABZ78073 Human bre
39	49	1.3	497	5	ABV58273 Human pro
40	49	1.3	882	2	AAI72173 Alzheimer
41	49	1.3	903	3	AAI72173 Alzheimer
42	49	1.3	1057	4	AAH33675 Human sec
43	49	1.3	1263	3	AAI79719 Human col
44	49	1.3	1373	3	AAI79719 Human col
45	49	1.3	1429	3	ABV94043 Breast ca
46	49	1.3	1676	4	AAH72649 Human cer
47	49	1.3	1702	5	ABV27467 Human pro
48	49	1.3	1702	5	ABV27724 Human pro
49	49	1.3	1702	5	ABV25281 Human pro
50	49	1.3	1702	5	ABV21893 Human pro
51	49	1.3	1933	4	AAI07369 Human DNA
52	49	1.3	1933	6	ABK92253 Prostate
53	49	1.3	1998	4	AAI16758 Human nov
54	49	1.3	1998	9	ADC22050 Human cDN
55	49	1.3	2038	7	ACD13422 Human DNA
56	49	1.3	2106	7	ABZ71965 Human ova
57	49	1.3	2143	6	ABQ54860 Human ORF
58	49	1.3	2543	3	AAI76644 Human CD2
59	49	1.3	2713	6	AAI43556 Human CD2
60	49	1.3	2713	6	AAI43556 Human CD2
61	49	1.3	2713	6	AAI43556 Human CD2
62	49	1.3	2713	6	AAI43556 Human CD2
63	49	1.3	2713	6	AAI43556 Human CD2
64	49	1.3	2851	6	ABN86559 Canine lo
65	49	1.3	2851	6	ABN86560 Canine CD
66	49	1.3	2944	7	ABT18794 Aspergill
67	49	1.3	3362	7	ABT20610 Aspergill
68	49	1.3	4236	6	ABQ92014 Human pol
69	49	1.3	4344	7	ABT18200 Aspergill
70	49	1.3	5362	7	ABT20014 Aspergill
71	48	1.3	295	4	AAH69364 Human cer
72	48	1.3	337	7	ABX40871 Bovine BS
73	48	1.3	401	8	ACH29107 Human adu
74	48	1.3	404	5	ABV44841 Human pro
75	48	1.3	418	4	AAI87343 Human pol
76	48	1.3	419	4	AAI88543 Human pol
77	48	1.3	546	4	AAH71167 Human cer
78	48	1.3	619	3	AAI26393 Human sec
79	48	1.3	865	6	ABN98811 Arabidops
80	48	1.3	903	4	AAI87708 Human pol
81	48	1.3	1136	3	AAI93117 Human sec
82	48	1.3	1467	3	AAI79042 Human sec
83	48	1.3	1510	6	ABK87557 CDNA enco
84	48	1.3	1544	3	AAI46483 CDNA enco
85	48	1.3	1837	6	ABK33555 CDNA enco
86	48	1.3	1837	7	ACA66860 CDNA enco
87	48	1.3	1837	7	ACD68612 Novel hum
88	48	1.3	1837	7	ACA68516 Novel hum
89	48	1.3	1837	8	ABT44245 Human PRO
90	48	1.3	1837	8	ABT44528 Human PRO
91	48	1.3	1837	8	ACD82195 Human sec
92	48	1.3	1837	8	ABT43901 Human men
93	48	1.3	1837	8	ABD83529 Novel hum
94	48	1.3	1837	8	ABD80635 Novel hum
95	48	1.3	1837	8	ABD73176 Novel hum
96	48	1.3	1837	8	ABD78258 Novel hum

97 48 1.3 1837 9 ADB84906
 98 48 1.3 1837 9 ADB78012
 99 48 1.3 1837 9 ADB87078
 100 48 1.3 1837 9 ADB84660
 101 48 1.3 1837 9 ADB83775
 102 48 1.3 1837 9 ADB72930
 103 48 1.3 1837 9 ADC36768
 104 48 1.3 1837 9 ADC21758
 105 48 1.3 1837 9 ADC49789
 106 48 1.3 1837 9 ADC48988
 107 48 1.3 1837 9 ADC49505
 108 48 1.3 1837 9 ADC47366
 109 48 1.3 1837 9 ADC47111
 110 48 1.3 1837 9 ADC77986
 111 48 1.3 1837 9 ADD06221
 112 48 1.3 1837 9 ADC77740
 113 48 1.3 1837 9 ADD50703
 114 48 1.3 1837 9 ADD50949
 115 48 1.3 1837 9 ADD50430
 116 48 1.3 1837 9 ADD50184
 117 48 1.3 1837 9 ADD51195
 118 48 1.3 1837 10 ADC48742
 119 48 1.3 1837 10 ADE20913
 120 48 1.3 1837 10 ADE05757
 121 48 1.3 1837 10 ADD74986
 122 48 1.3 1837 10 ADD75732
 123 48 1.3 1837 10 ADD84964
 124 48 1.3 1837 10 ADD86790
 125 48 1.3 1837 10 ADE20667
 126 48 1.3 1837 10 ADE38964
 127 48 1.3 1837 10 ADE05511
 128 48 1.3 1837 10 ADD73496
 129 48 1.3 1837 10 ADD78336
 130 48 1.3 1837 10 ADE21159
 131 48 1.3 1837 10 ADD77274
 132 48 1.3 1837 10 ADE20421
 133 48 1.3 1837 10 ADD75486
 134 48 1.3 1837 10 ADD74002
 135 48 1.3 1837 10 ADD74248
 136 48 1.3 1837 10 ADD75978
 137 48 1.3 1837 10 ADD85470
 138 48 1.3 1837 10 ADE05019
 139 48 1.3 1837 10 ADD75232
 140 48 1.3 1837 10 ADD76776
 141 48 1.3 1837 10 ADD86544
 142 48 1.3 1837 10 ADD78012
 143 48 1.3 1837 10 ADD77520
 144 48 1.3 1837 10 ADD77766
 145 48 1.3 1837 10 ADD85224
 146 48 1.3 1837 10 ADD73756
 147 48 1.3 1837 10 ADD74494
 148 48 1.3 1837 10 ADD77022
 149 48 1.3 1837 10 ADD85716
 150 48 1.3 1837 10 ADE05265

ALIGNMENTS

RESULT 1
 AEN81319 standard; cDNA; 3762 BP.
 ID AEN81319
 AC AEN81319;
 XX
 DT 30-AUG-2002 (first entry)
 XX
 DE Human mast cell related gene MCI SEQ ID NO 1.
 XX
 KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
 KW vasotrophic; dermatological; allergic; hypersensitivity; rhinitis; asthma;
 KW gene; ss.
 XX

OS Homo sapiens.
 Key Location/Qualifiers
 CDS 25..432
 /tag= a
 /product= "MCI"
 WO2002046389-A2.
 13-JUN-2002.
 07-DEC-2001; 2001WO-US046180.
 08-DEC-2000; 2000US-0251835P.
 14-MAR-2001; 2001US-0275479P.
 28-MAR-2001; 2001US-0279115P.
 02-APR-2001; 2001US-0280143P.
 (UNIO) UCB SA.
 Nocka K, Pirozzi G, Einstein R;
 WPI; 2002-508560/54.
 P-PSDB; ABB77569.
 Novel isolated nucleic acids that are differentially expressed in mast cells in patients with allergic hypersensitivity, encoding proteins associated with mast cell regranulation and allergic hypersensitivity.
 Claim 1; Page 95-97; 119pp; English.
 The invention relates to isolated nucleic acid (AEN81319-AEN81324), corresponding to genes differentially expressed in mast cells following activation or in patients with allergic hypersensitivity disease, (I) that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of (II) if at least 6 amino acids. (II) is useful for identifying binding partners. (I) or (II) is useful for diagnosing or treating a disease state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma, urticaria or atopic dermatitis or mastocytosis) in a subject which involves determining the level of expression of (I) or (II). A computer system, comprising a database containing information identifying the expression level in a tissue or at least one mast cell of (I), is useful for presenting information to identify the relative expression level of (I). (II) is used as a marker to detect, diagnose or identify an allergic response in a patient. The protein can also serve as a target that modulate gene expression or activity and as an antigen to raise polyclonal or monoclonal antibodies. (II) is useful for identifying agents that modulate expression of the protein or agents, such as agonists or antagonists. The agonists or antagonists are useful for modulating biological activity and function of (II) and thus are useful for alleviating disease conditions such as allergic hypersensitivity, seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
 Sequence 3762 BP; 1220 A; 672 C; 580 G; 1190 T; 0 U; 0 Other;
 Query Match 100.0%; Score 3762; DB 6; Length 3762;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3762; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGAAACCGAGTCACTGTGAAAGATGGAAATATCTCTCGGAAACTCAGTTGCTG 60
 DB 1 GAGAAACCGAGTCACTGTGAAAGATGGAAATATCTCTCGGAAACTCAGTTGCTG 60
 QY 61 GGAGAGATCAAAAGAACCCCAAGAAAGAAACCCAGATGAGGAAAGAAACGGCAGAA 120
 DB 61 GGAGAGATCAAAAGAACCCCAAGAAAGAAACCCAGATGAGGAAAGAAACGGCAGAA 120
 QY 121 ATGACTACATTGAAAGAAACTCAAGATCAAGTAGAAGAAAGCAAGAGATTTCATCC 180
 DB 121 ATGACTACATTGAAAGAAACTTCAGATCAAGTAGAAGAAAGCAAGAGATTTCATCC 180
 QY 181 ACTTCTAATCAGGAAACGAGAAATGGCAGTGGTCTTGAAGAAAGTGTGTACATGTCTATT 240

181 ACTTCTAATCAGGAAACGAGAAATGCGAGTGGTTCTGAAGAGTGTCTACACTGTCATT 240
182
241 AATCACATCCCCCATCAGAGATCTCCCTGAGCTCCAAATGATGATGCTATGAGACATT 300
242
241 AATCACATCCCCCATCAGAGATCTCCCTGAGCTCCAAATGATGCTATGAGACATT 300
243
301 GACTCCCTCAAGGAAAGTGAGACAGTGTAGAGAAAGGTGAGAGACAGAAATATGCCCTT 360
302
301 GACTCCCTCAAGGAAAGTGAGACAGTGTAGAGAAAGGTGAGAGACAGAAATATGCCCTT 360
303
361 CTTAGGACTTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATGATGAGTGTG 420
362
361 CTTAGGACTTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATGATGAGTGTG 420
363
421 TTTCCACACATAAATCCTCAAGCTGCTTTATCACCCTTCCAGCAATGAAGACATGAGAA 480
422
421 TTTCCACACATAAATCCTCAAGCTGCTTTATCACCCTTCCAGCAATGAAGACATGAGAA 480
423
481 TAGCAGACTCTGGGAAAGTTGTTACCCCTGAGCAGTGCATGAACAATTCCTTCTGGCTA 540
482
481 TAGCAGACTCTGGGAAAGTTGTTACCCCTGAGCAGTGCATGAACAATTCCTTCTGGCTA 540
483
541 AAGTTTGAAGATATATCTTATATATATATCTTATGAGCACTCTGATATGAGCATCTCG 600
542
541 AAGTTTGAAGATATATCTTATATATATATCTTATGAGCACTCTGATATGAGCATCTCG 600
543
601 TGGCTTAGGTGAATCATAGAAATGACACAATGACCTAAATATTTCTATGCTTTTTC 660
602
601 TGGCTTAGGTGAATCATAGAAATGACACAATGACCTAAATATTTCTATGCTTTTTC 660
603
661 TTGTTAAGTTTGAAGCATGGAGTGTATATAAATAAATCTTCTAGGACATATATGTA 720
662
661 TTGTTAAGTTTGAAGCATGGAGTGTATATAAATAAATCTTCTAGGACATATATGTA 720
663
721 TGAATAATAATTTCTAATCCCTCTGACTAATGAGTGAACCTCTCTAGGCAAAAGAGA 780
722
721 TGAATAATAATTTCTAATCCCTCTGACTAATGAGTGAACCTCTCTAGGCAAAAGAGA 780
723
781 CCTCAGATGACCTGAAGACTGATATCTGGCCATGATAGAGAGGAGTGAACACACC 840
782
781 CCTCAGATGACCTGAAGACTGATATCTGGCCATGATAGAGAGGAGTGAACACACC 840
783
841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGAACAGGATGAGTCAATAGACT 900
842
841 TTGTTATACCCCTTCCCTTTTGGAGTTTATGCAAGTGAACAGGATGAGTCAATAGACT 900
843
901 GATGAATATAGACTGATTTGGCAATAGAGTCCCAATTCCAACTGACTGCTGTGTAGAT 960
902
901 GATGAATATAGACTGATTTGGCAATAGAGTCCCAATTCCAACTGACTGCTGTGTAGAT 960
903
961 CACACACTGCTGAGGATTCATCTATGAGACTTTGCTACATATACAGAGACCTTGGTT 1020
962
961 CACACACTGCTGAGGATTCATCTATGAGACTTTGCTACATATACAGAGACCTTGGTT 1020
963
1021 TCCACACCCCTTTATTTAGCTAAAGATCTCTTTCTACTGACTTTAAAGTCTTTAGA 1080
1022
1021 TCCACACCCCTTTATTTAGCTAAAGATCTTTTCTACTGACTTTAAAGTCTTTAGA 1080
1023
1081 CAAGCTTAACTCTTTCAACCAATGCGAATCAGACAAATTTGAATCTACCTATGACCT 1140
1082
1081 CAAGCTTAACTCTTTCAACCAATGCGAATCAGACAAATTTGAATCTACCTATGACCT 1140
1083
1141 GTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTTAAAGCTGAACCGATGTCACCTTCCA 1200
1142
1141 GTAAGCTCTCTCTGCTTCAAGATCTTGGCTCTTTTAAAGCTGAACCGATGTCACCTTCCA 1200
1143
1201 TTTTAACTGATTTAGTCTTCTGCTTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
1202
1201 TTTTAACTGATTTAGTCTTCTGCTTGAATCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1260
1203
1261 GTGACCTGACCACTCAGGCACACTTCTCAGGACCTCTCTGAGAGTGTATCCAGGACCAT 1320
1262
1261 GTGACCTGACCACTCAGGCACACTTCTCAGGACCTCTCTGAGAGTGTATCCAGGACCAT 1320

1321 GGTAACTCATCTTGGCTCAGAAATCAACTCTTTTAAATATTTTACAGAAATTTGGGTTTGG 1380
1322
1321 GGTAACTCATCTTGGCTCAGAAATCAACTCTTTTAAATATTTTACAGAAATTTGGGTTTGG 1380
1323
1381 TTACCAATAAGTCTCCCAAAATATATGTCCAAAGAAATCTTCAATTTCCAAAGCTGCTCACC 1440
1382
1381 TTACCAATAAGTCTCCCAAAATATATGTCCAAAGAAATCTTCAATTTCCAAAGCTGCTCACC 1440
1383
1441 AATTTCAAAATGCCAAATCTCCCATCAATTTCAATTTCTATTTCTATTTCTGAGGTGTAATCTA 1500
1442
1441 AATTTCAAAATGCCAAATCTCCCATCAATTTCAATTTCTATTTCTATTTCTGAGGTGTAATCTA 1500
1443
1501 CTCATAAATCTGTGTAGACCAAGTGCAGACACCTTTTGTAACTGACATTTACTTCAAT 1560
1502
1501 CTCATAAATCTGTGTAGACCAAGTGCAGACACCTTTTGTAACTGACATTTACTTCAAT 1560
1503
1561 TTTTCTTTTCTATGCTACTGATATTTTGCATATAAATCTTGCAGTAAATAGTTCAAAAAT 1620
1562
1561 TTTTCTTTTCTATGCTACTGATATTTTGCATATAAATCTTGCAGTAAATAGTTCAAAAAT 1620
1563
1621 TAAATAGTTTGTGACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680
1622
1621 TAAATAGTTTGTGACATTTGGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAA 1680
1623
1681 AAGATGAATGAAGCATATATAATTTGTCATATAAATTTGCAATTTCTAGTCAACAGAAATC 1740
1682
1681 AAGATGAATGAAGCATATATAATTTGTCATATAAATTTGCAATTTCTAGTCAACAGAAATC 1740
1683
1741 GAAGGATTTCTGTTCAATATTTAGTAAATAATTTGAAATTAATCTGCTTATATTTTGT 1800
1742
1741 GAAGGATTTCTGTTCAATATTTAGTAAATAATTTGAAATTAATCTGCTTATATTTTGT 1800
1743
1801 GCACACACTAGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGGTGATGTAGT 1860
1802
1801 GCACACACTAGTTAAATTTAACTGTGACTAGTATCTCTACCGAAGGTGATGTAGT 1860
1803
1861 TTCTGCTTTTAAATTTCAAGCAAACTGAAATTAATCCATCTAATTTATGCTTTCTTTCCC 1920
1862
1861 TTCTGCTTTTAAATTTCAAGCAAACTGAAATTAATCCATCTAATTTATGCTTTCTTTCCC 1920
1863
1921 AAGAAAGTTTAAATGATATGCCAGTCTTCTAAATTTGGAGACAAAGCTTAAATGACAA 1980
1922
1921 AAGAAAGTTTAAATGATATGCCAGTCTTCTAAATTTGGAGACAAAGCTTAAATGACAA 1980
1923
1981 TGCATTCATATATATTTTGTATAGTTACAGTATACAGTTCAGTATCCCTTAGATG 2040
1982
1981 TGCATTCATATATATTTTGTATAGTTACAGTATACAGTTCAGTATCCCTTAGATG 2040
1983
2041 AGATGCTTTGGGACCAAGAGTGTTTTGGATTTTCAAGTTTATTTTGGATTTTGGATATTT 2100
2042
2041 AGATGCTTTGGGACCAAGAGTGTTTTGGATTTTCAAGTTTATTTTGGATTTTGGATATTT 2100
2043
2101 CCATACATATATGAGAGAGTGTGAAATGGATTCAGTCTAATTCATATAAATTCACCTTA 2160
2102
2101 CCATACATATATGAGAGAGTGTGAAATGGATTCAGTCTAATTCATATAAATTCACCTTA 2160
2103
2161 TGTGTTGATATACCTCTATCTGAATAGCTCTGAAGTAAATTTTATACATATTTTAAATAA 2220
2162
2161 TGTGTTGATATACCTCTATCTGAATAGCTCTGAAGTAAATTTTATACATATTTTAAATAA 2220
2163
2221 TTTTATGCTGAAACAGAGTGTGGCACAATTTGGACCATCAGAAAGCAGAGTGTACCTAT 2280
2222
2221 TTTTATGCTGAAACAGAGTGTGGCACAATTTGGACCATCAGAAAGCAGAGTGTACCTAT 2280
2223
2281 TTTCAAGTCAAGTGTCTCAAAAGTTTTCAGATGTTAGCTGCTGATGAGTTCATGCCAGTGA 2340
2282
2281 TTTCAAGTCAAGTGTCTCAAAAGTTTTCAGATGTTAGCTGCTGATGAGTTCATGCCAGTGA 2340
2283
2341 TCCGAGTACTTTTGGAGCCCAAGACAGGTGATCTTTGAGCCCAAGGAGTTTGGGCCAG 2400
2342
2341 TCCGAGTACTTTTGGAGCCCAAGACAGGTGATCTTTGAGCCCAAGGAGTTTGGGCCAG 2400

QY 2401 ACTGCACAAACAGGTAGAGACCTCGTTCTTCTACAAATTAATAAAATTTAGCCAGGTGTGGT 2460
DB 2401 ACTGCACAAACAGGTAGAGACCTCGTTCTTCTACAAATTAATAAAATTTAGCCAGGTGTGGT 2460
QY 2461 GGTGCACACCTGTAGTCCCGAGTACTCAGGAGGCTGAGGTAGTAGGATTTGTTGAGACTG 2520
DB 2461 GGTGCACACCTGTAGTCCCGAGTACTCAGGAGGCTGAGGTAGTAGGATTTGTTGAGACTG 2520
QY 2521 GGAGGTTGAGGCTGAACGTGAGCCAGGATCTTTGCCACCACATTCAGCTTGGGCAACAGAG 2580
DB 2521 GGAGGTTGAGGCTGAACGTGAGCCAGGATCTTTGCCACCACATTCAGCTTGGGCAACAGAG 2580
QY 2581 TGAGACCTGTCTCAAAAAAAGATTTTCAGATTTTGGAGCATTTTCAGATCTTCAG 2640
DB 2581 TGAGACCTGTCTCAAAAAAAGATTTTCAGATTTTGGAGCATTTTCAGATCTTCAG 2640
QY 2641 ATTAGGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTGCAAGCATTTATCATAGGTG 2700
DB 2641 ATTAGGGATTTTCAACCTGTACTGACCTTTTGTAGTCATTGCAAGCATTTATCATAGGTG 2700
QY 2701 GACTCCAGATTAACCTTTGCTGTATACACATTTTGGCTCTCTATTCAAGCAATTCCTTAT 2760
DB 2701 GACTCCAGATTAACCTTTGCTGTATACACATTTTGGCTCTCTATTCAAGCAATTCCTTAT 2760
QY 2761 GCCCTCTTGTGGTGAATTTAATGTGCGAAGGAAACAAATAGAAATTTTGCATTTCTAGA 2820
DB 2761 GCCCTCTTGTGGTGAATTTAATGTGCGAAGGAAACAAATAGAAATTTTGCATTTCTAGA 2820
QY 2821 AAAGTCATCTGTCAAAATATGTGAGTCTGTAGATATTAGCCAAATTTAGGAAATGAC 2880
DB 2821 AAAGTCATCTGTCAAAATATGTGAGTCTGTAGATATTAGCCAAATTTAGGAAATGAC 2880
QY 2881 AAAATTTTCTTTCCTGCTGCTTGTAGTCTGTTTATGATATAATACCTTATTTGT 2940
DB 2881 AAAATTTTCTTTCCTGCTGCTTGTAGTCTGTTTATGATATAATACCTTATTTGT 2940
QY 2941 AATAAAATTAATTTTAAATTTGAGTAACAAATCTGGAATTTATCAGAGAAGGGCAAGCAATA 3000
DB 2941 AATAAAATTAATTTTAAATTTGAGTAACAAATCTGGAATTTATCAGAGAAGGGCAAGCAATA 3000
QY 3001 GGTAAATAACAGTATTTGATTTGAGAGGACCTTGAAATCCAGAGCATCAATGCTT 3060
DB 3001 GGTAAATAACAGTATTTGATTTGAGAGGACCTTGAAATCCAGAGCATCAATGCTT 3060
QY 3061 CTGGTGGTTTCCACCAATACCCACAGACAGATGCTTAATCTTTCCGAGATCTAGTTTTTCAG 3120
DB 3061 CTGGTGGTTTCCACCAATACCCACAGACAGATGCTTAATCTTTCCGAGATCTAGTTTTTCAG 3120
QY 3121 CAAAGCAGATTTAAGAAATGTAATCTTATGTTGTTTGAAGAACATAGAAATCAAT 3180
DB 3121 CAAAGCAGATTTAAGAAATGTAATCTTATGTTGTTTGAAGAACATAGAAATCAAT 3180
QY 3181 GCTGTATAAGTGTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTTTATGATATAAAT 3240
DB 3181 GCTGTATAAGTGTCTTTTAACTGTAAATTTTGTGAAGCTTATCTTTTATGATATAAAT 3240
QY 3241 ATTGAAACATTTTCAATTTTATTTTAAATCAAGTTTACTCAAGTGTGATTAATAC 3300
DB 3241 ATTGAAACATTTTCAATTTTATTTTAAATCAAGTTTACTCAAGTGTGATTAATAC 3300
QY 3301 AAGAAATGTAACCACTGTAGGGTAGAGTTTAAAGATTTTGTCAATTTGATTCACCCA 3360
DB 3301 AAGAAATGTAACCACTGTAGGGTAGAGTTTAAAGATTTTGTCAATTTGATTCACCCA 3360
QY 3361 TGTAGTCACTCTCTTATGAAGAGACAGAACAGTACATCTCTCCCAAGAGTTCCACAGTG 3420
DB 3361 TGTAGTCACTCTCTTATGAAGAGACAGAACAGTACATCTCTCCCAAGAGTTCCACAGTG 3420
QY 3421 CTCCTTTTCCCTGAGTTTCCAGCTCTGCGCAACCAATGATCTGCTCGTATATTATAA 3480
DB 3421 CTCCTTTTCCCTGAGTTTCCAGCTCTGCGCAACCAATGATCTGCTCGTATATTATAA 3480
QY 3481 CTGTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTTATTTTGTGTGTGAAGGCTT 3540

DB 3481 CTGTTCTAGATATTTGTAGCAATGTACCTTTTCCATATTTTATTTGTGTGTGAAGGCTT 3540
QY 3541 CTTTGTAGTCATTAATAATTTTGTAGATTTTCATCTATGTTTAAATGTTCTATCAGTAGTGT 3600
DB 3541 CTTTGTAGTCATTAATAATTTTGTAGATTTTCATCTATGTTTAAATGTTCTATCAGTAGTGT 3600
QY 3601 ACATCTTACTTCTCAGCATATCACCATATAGATATATCTAATTTTGTAAATCTAATCA 3660
DB 3601 ACATCTTACTTCTCAGCATATCACCATATAGATATATCTAATTTTGTAAATCTAATCA 3660
QY 3661 CTGATGGATATCTAGATATTTAAGTTTTTGCATTTATGAATTAAGTGGCTATTAATGAA 3720
DB 3661 CTGATGGATATCTAGATATTTAAGTTTTTGCATTTATGAATTAAGTGGCTATTAATGAA 3720
QY 3721 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 3762
DB 3721 AAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTTAAAAAATTTT 3762

RESULT 2
ABV20298
ID ABV20298 standard; cDNA; 1128 BP.
XX
AC ABV20298;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 20289.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN W0200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
PS WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3320-3321; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

QY 2393 GAGCCGACGCTGCACACACACAGTGAGACCTCGTTTCT 2429
 DB 958 GAGCCGACGCTGCACACACACAGTGAGACCTCGTTTCT 994

RESULT 6

ID ABV20806 standard; cDNA; 1128 BP.

AC ABV20806;

DT 13-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 20797.

KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

FN WO200160860-A2.

PD 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 3412-3413; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV0010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the efficacy of a test compound to inhibit prostate cancer in a patient; (c) assessing the progression of prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 19.5%; Score 733; DB 5; Length 1128;

Best Local Similarity 99.6%; Pred. No. 3.5e-244;

Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1493 GTAATCTACTAATAAAGTGTATAGACAGTGACAGACCTTTGCTAACTGACATTT 1552

DB 58 GTAATCTACTAATAAAGTGTATAGACAGTGACAGACCTTTGCTAACTGACATTT 117

QY 1553 ACTTCAATTTTCTTTTCTATGTACTGATATTTTGCATATAAAGTGTAGTAAAGT 1612

DB 118 ACTTCAATTTTCTTTTCTATGTACTGATATTTTGCATATAAAGTGTAGTAAAGT 177

QY 1613 TCAAAAATTAATAGTATTTTGCATTTTCTCGAAGAGAGAAATGAAAGTGTCAAA 1672

DB 178 TCAAAAATTAATAGTATTTTGCATTTTCTCGAAGAGAGAAATGAAAGTGTCAAA 237
 QY 1673 AATAAAAAAGATGAAATGAAGCATATATAATGTCAATTTTTCATTTTCTAGTCAAC 1732
 DB 238 AATAAAAAAGATGAAATGAAGCATATATAATGTCAATTTTTCATTTTCTAGTCAAC 297
 QY 1733 AGAAGATCGAAGATCTGTTCAAAATATAGTAAAAATGAAAAATAACTTTGTGCTTATA 1792
 DB 298 AGAAGATCGAAGATCTGTTCAAAATATAGTAAAAATGAAAAATAACTTTGTGCTTATA 357
 QY 1793 TTTTGTTCGCAACACACTAGTTAAATTTAACCTGTGACTAGTTATCTCTACCGAAGGTGA 1852
 DB 358 TTTTGTTCGCAACACACTAGTTAAATTTAACCTGTGACTAGTTATCTCTACCGAAGGTGA 417
 QY 1853 TGTGTAGTTTCTGTGTTTAAATTCGAAGCAAACTCGAAAAATTAATCCATCTAATATGCTT 1912
 DB 418 TGTGTAGTTTCTGTGTTTAAATTCGAAGCAAACTCGAAAAATTAATCCATCTAATATGCTT 477
 QY 1913 TCTTTCCCAAGAGTTTAAATTAATGATGCGCAGCTTCTTAATTTGCGAGACAAAGCCCTTA 1972
 DB 478 TCTTTCCCAAGAGTTTAAATTAATGATGCGCAGCTTCTTAATTTGCGAGACAAAGCCCTTA 537
 QY 1973 ATTGACATGCAATTCATTATATATTTTGTATAGTTTACAGTATACGAGTTGAGTATCC 2032
 DB 538 ATTGACATGCAATTCATTATATATTTTGTATAGTTTACAGTATACGAGTTGAGTATCC 597
 QY 2033 CTTAGATGAGATGCTTGGGACGAGAGTTTGGGATTTTGGATTTTATTTTGGATTTTG 2092
 DB 598 CTTAGATGAGATGCTTGGGACGAGAGTTTGGGATTTTGGATTTTATTTTGGATTTTG 657
 QY 2093 GAATATTTCCATACATATATAATGAGAGAGTTGGAAAAATGGATTCAGTCTAATCAATAA 2152
 DB 658 GAATATTTCCATACATATAATGAGAGAGTTGGAAAAATGGATTCAGTCTAATCAATAA 717
 QY 2153 TTCACTTATGTTGATATACACTTATCTGAATAGCTGAGAGTTTATTTTATCAATATT 2212
 DB 718 TGCACTTATGTTGATATACACTTATCTGAATAGCTGAGAGTTTATTTTATCAATATT 777
 QY 2213 TTAATAATTTTATGCTGAAACAGAGTTTGGGCACTTGGACCATCGAAGAGAGAGT 2272
 DB 778 TTAATAATTTTATGCTGAAACAGAGTTTGGGCACTTGGACCATCGAAGAGAGAGT 837
 QY 2273 GTCATATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGTATCGAGTTCA 2332
 DB 838 GTCATATTTCAAGTCAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGTATCGAGTTCA 897
 QY 2333 GCCAGTGATCCGAGTACTTTGGGAGCCCAAGACAGTGTGATCTCTTGGCCCGAGAGTTT 2392
 DB 898 GCCAGTGATCCGAGTACTTTGGGAGCCCAAGACAGTGTGATCTCTTGGCCCGAGAGTTT 957
 QY 2393 GAGGCCAGACTGCACAAACACAGTGAGACCTCGTTTCT 2429
 DB 958 GAGGCCAGACTGCACAAACACAGTGAGACCTCGTTTCT 994

RESULT 7

ABV20439

ID ABV20439 standard; cDNA; 1128 BP.

XX AC ABV20439;

XX DT 13-SEP-2002 (first entry)

XX DE Human prostate expression marker cDNA 20430.

XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.

XX OS Homo sapiens.

XX FN WO200160860-A2.

XX XX

PD 23-AUG-2001.
 XX 20-FEB-2001; 2001WO-US005171.
 XX 17-FEB-2000; 2000US-0183319P.
 XX 16-MAR-2000; 2000US-0189862P.
 XX 25-MAY-2000; 2000US-0207454P.
 XX 09-JUN-2000; 2000US-0211314P.
 XX 18-JUL-2000; 2000US-0219007P.
 XX 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 PT Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 3347; 11750pp; English.
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;
 Query Match 19.5%; Score 733; DB 5; Length 1128;
 Best Local Similarity 99.6%; Pred. No. 3.5e-244;
 Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
 QY 1493 GTAATCTACTCAATAAAGCTGTGAAGACAGACAGACCTTTGTAACCTGACATTT 1552
 DB 58 GTAATCTACTCAATAAAGCTGTGAAGACAGACAGACCTTTGTAACCTGACATTT 117
 QY 1553 ACTTCAATTTCTTTTCTATGCTGATATTTTGCATATTAACCTGCAAGTAAGT 1612
 DB 118 ACTTCAATTTCTTTTCTATGCTGATATTTTGCATATTAACCTGCAAGTAAGT 177
 QY 1613 TCAAAAATTAATAGTTTTTGTGACATGGCTTTTCTGAGAAGAGAAATGAAAGTGTCAA 1672
 DB 178 TCAAAAATTAATAGTTTTTGTGACATGGCTTTTCTGAGAAGAGAAATGAAAGTGTCAA 237
 QY 1673 AATAAAAAGATGAATGAGCATATATATTTGTCATTTTTTCAATTTTCTAGTCAAC 1732
 DB 238 AATAAAAAGATGAATGAGCATATATATTTGTCATTTTTTCAATTTTCTAGTCAAC 297
 QY 1733 AGAAGATCGAAGGATTCGTTTCAATATAGTAAATAATGAAATAAATGCTGCTTATA 1792
 DB 298 AGAAGATCGAAGGATTCGTTTCAATATAGTAAATAATGAAATAAATGCTGCTTATA 357
 QY 1793 TTTGTTTGTGACACACATAGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGA 1852
 DB 358 TTTGTTTGTGACACACATAGTTAATTTAACTGTGACTAGTTATCTCTACCGAAGTGA 417
 QY 1853 TGTGTAGTTCTGTTTAAATAATCAAGCAACTGGAATAATCAATCTAATATATGCTT 1912
 DB 418 TGTGTAGTTCTGTTTAAATAATCAAGCAACTGGAATAATCAATCTAATATATGCTT 477
 QY 1913 TCTTTTCCCAAGAGTTTTTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAGCCTTA 1972
 DB 478 TCTTTTCCCAAGAGTTTTTTTAAATGATATGCGAGCTTCTTAATTTGGAGACAAAGCCTTA 537

QY 1973 ATTGCAATGCATTCATTATATATTTTTTGTATAGTTTACAGTATACAGTTGAGTATCC 2032
 DB 538 ATTGCAATGCATTCATTATATATTTTTTGTATAGTTTACAGTATACAGTTGAGTATCC 597
 QY 2033 CTTAGATGAGATGCTTGGGACAGAGTGTGTTGGATTTTCAGATTTATTTTGGATTTG 2092
 DB 598 CTTAGATGAGATGCTTGGGACAGAGTGTGTTGGATTTTCAGATTTATTTTGGATTTG 657
 QY 2093 GAATATTTCCATACATATAATGAGAGAGTTGAAAATGGGATTTCAAGTCTAATCATATAAA 2152
 DB 658 GAATATTTCCATACATATAATGAGAGAGTTGTAATGGGATTTCAAGTCTAATCATATAAA 717
 QY 2153 TTCACCTATGTTGATATACACCTTATCTGAATAGCCTGAGGTAATTTTATACATATTT 2212
 DB 718 TGCACCTATGTTGATATACACCTTATCTGAATAGCCTGAGGTAATTTTATACATATTT 777
 QY 2213 TTAATAATTTTATGCTTGAACACAGATTGGGACCAATTTGGACCAATCAGAAAGCAGAGT 2272
 DB 778 TTAATAATTTTATGCTTGAACACAGATTGGGACCAATTTGGACCAATCAGAAAGCAGAGT 837
 QY 2273 GTCACCTATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTGATGCAATTCAT 2332
 DB 838 GTCACCTATTTCAAGTCAGTGTCTCAAAAAGTTTTCAGATGTTAAGCTGGTGATGCAATTCAT 897
 QY 2333 GCACGTGATCCGAGTACTTTGGGAGCCAGACAGAGTGTGATCTCTTGAGCCCGAGGATTT 2392
 DB 898 GCACGTGATCCGAGTACTTTGGGAGCCAGACAGAGTGTGATCTCTTGAGCCCGAGGATTT 957
 QY 2393 GAGGCCAGACTCTCACAACACACAGTGTGAGACTCGTTTCT 2429
 DB 958 GAGGCCAGACTCTCACAACACACAGTGTGAGACTCGTTTCT 994
 RESULT 8
 ABV26275
 ID ABV26275 standard; cDNA; 1128 BP.
 XX AC ABV26275;
 XX AC ABV26275;
 DT 16-SEP-2002 (first entry)
 XX Human prostate expression marker cDNA 26266.
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 OS Homo sapiens.
 XX MO200160860-A2.
 XX 23-AUG-2001.
 XX 20-FEB-2001; 2001WO-US005171.
 XX 17-FEB-2000; 2000US-0183319P.
 XX 16-MAR-2000; 2000US-0189862P.
 XX 25-MAY-2000; 2000US-0207454P.
 XX 09-JUN-2000; 2000US-0211314P.
 XX 18-JUL-2000; 2000US-0219007P.
 XX 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 5313; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 19.5%; Score 733; DB 5; Length 1128;
Best Local Similarity 99.6%; Pred. No. 3.5e-244;
Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1493 GTAATCTACTCAATAAAGCTGTAGACAGTGCAGTGCAGCAGCCTTGTCAACTGACATTT 1552
Db |||||
QY 58 GTAATCTACTCAATAAAGCTGTAGACAGTGCAGTGCAGCAGCCTTGTCAACTGACATTT 117
Db |||||
QY 1553 ACTTCAATTTTCTTTTCTATGTACTGATATTTTGCATATATACTTGCAGTAATAGT 1612
Db |||||
QY 1613 TCAAAATTAATAGTTTGTGACATTTGCTTCTGAGAGAGAAATGAAAGTGCACAA 1672
Db |||||
QY 178 TCAAAATTAATAGTTTGTGACATTTGCTTCTGAGAGAGAAATGAAAGTGCACAA 237
Db |||||
QY 1673 AATAAATAAGATGAAATGAACATATATAATGTCAATTTTCAATTTCTAGTCAAC 1732
Db |||||
QY 238 AATAAATAAGATGAAATGAACATATATAATGTCAATTTTCAATTTCTAGTCAAC 297
Db |||||
QY 1733 AGAGATCGAAGATTCGTTCAATATATAGTAAATTAATGAAATTAATGCTTATA 1792
Db |||||
QY 298 AGAGATCGAAGATTCGTTCAATATATAGTAAATTAATGAAATTAATGCTTATA 357
Db |||||
QY 1793 TTTTGTGTCACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGA 1852
Db |||||
QY 358 TTTTGTGTCACACACTAGTTAAATTAACCTGTGACTAGTTATCTCTACCGAAGGTGGA 417
Db |||||
QY 1853 TGTAGTATTCGTTTAAATTAATCAAGCAACTGGAATAATATCCATTAATATGCTT 1912
Db |||||
QY 418 TGTAGTATTCGTTTAAATTAATCAAGCAACTGGAATAATATCCATTAATATGCTT 477
Db |||||
QY 1913 TCTTTCCCAAGATTTTATATGATATGCCAGCTTCCTAATTTGGAGACAAAGCCTTA 1972
Db |||||
QY 478 TCTTTCCCAAGATTTTATATGATATGCCAGCTTCCTAATTTGGAGACAAAGCCTTA 537
Db |||||
QY 1973 ATGCAATGCAATTCATATATATTTTGTATAGTTTACAGTATACAGTGTGATATCC 2032
Db |||||
QY 538 ATGCAATGCAATTCATATATATTTTGTATAGTTTACAGTATACAGTGTGATATCC 597
Db |||||
QY 2033 CTTAGTGCAGTCTGGGACGAGGTGTTTGGATTCAGATTTATTTTGTGATTTG 2092
Db |||||
QY 598 CTTAGTGCAGTCTGGGACGAGGTGTTTGGATTCAGATTTATTTTGTGATTTG 657
Db |||||
QY 2093 GAATATTTCCATATATATGAGAGGTGGAATAATGGATTCAGTCTAATCATAAA 2152
Db |||||
QY 658 GAATATTTCCATATATGAGAGGTGGAATAATGGATTCAGTCTAATCATAAA 717
Db |||||
QY 2153 TTCACCTATGTTGATATACACCTTATCTGATAGCTGAAGGTAAATTTTATCAATTT 2212
Db |||||
QY 718 TGCACCTATGTTGATATACACCTTATCTGATAGCTGAAGGTAAATTTTATCAATTT 777
Db |||||
QY 2213 TTAATAATTTTATGCTGAAACAGAGTTTGGCAGATTTGGCAGATTCAGAAAGCAGAAGT 2272
Db |||||
QY 778 TTAATAATTTTATGCTGAAACAGAGTTTGGCAGATTTGGCAGATTCAGAAAGCAGAAGT 837
Db |||||
QY 2273 GTCACATTTCAAGTCAGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGTGATTCAT 2332
Db |||||

Db 838 GTCACATTTTCAAGTCAGTCTCAAAAAGTTTCAGATGTTAAGCTGGTGTGATTCAT 897
QY 2333 GCCAGTGCATCCAGTACTTTGGAGAGCCAAAGACAGTGTGATCTCTTTGAGCCAGAGTTT 2392
Db 898 GCCAGTGCATCCAGTACTTTGGAGAGCCAAAGACAGTGTGATCTCTTTGAGCCAGAGTTT 957
QY 2393 GAGGCCAGACTGCACAAACACAGTGCAGTGCAGTCTCTTTCT 2429
Db 958 GAGGCCAGACTGCACAAACACAGTGCAGTGCAGTCTCTTTCT 994

RESULT 9
ABV26653
ID ABV26653 standard; cDNA; 1128 BP.
XX AC ABV26653;
XX XX
DT 16-SEP-2002 (first entry)
XX XX
DE Human prostate expression marker cDNA 26644.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
OS
XX WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
XX WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 5383; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 1128 BP; 354 A; 180 C; 216 G; 374 T; 0 U; 4 Other;

Query Match 19.5%; Score 733; DB 5; Length 1128;
Best Local Similarity 99.6%; Pred. No. 3.5e-244;
Matches 933; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1493 GTAATCTACTCAATAAAGCTGTAGACAGTGCAGTGCAGCAGCCTTGTCAACTGACATTT 1552
Db |||||

Db 58 GTAATCTACTCAATAAACTGTGAAGACCACTGACAGACCCCTTGCTAACCTGCACATTT 117
Qy 1553 ACTTCATTTTCTTTTCTATGCTAGTGGATATTTTGGATATAAACTTGCAGTATAGT 1612
Db 118 ACTTCATTTTCTTTTCTATGCTAGTGGATATTTTGGATATAAACTTGCAGTATAGT 177
Qy 1613 TCAGAAATTAATAGTTTTGACATTCGCTTTTCTGAGAGAGAAATGGAAGTGTCAACA 1672
Db 178 TCAGAAATTAATAGTTTTGACATTCGCTTTTCTGAGAGAGAAATGGAAGTGTCAACA 237
Qy 1673 AATAAATAAGATGAATGAAGCATATATAATGTGCAATTTTCAATTTTCTAGTCAAC 1732
Db 238 AATAAATAAGATGAATGAAGCATATATAATGTGCAATTTTCAATTTTCTAGTCAAC 297
Qy 1733 AGAATTCGAAGGATTCGTTCCTCAATATAGTAAAAATGGAATTAACCTTGCTTATA 1792
Db 298 AGAATTCGAAGGATTCGTTCCTCAATATAGTAAAAATGGAATTAACCTTGCTTATA 357
Qy 1793 TTTTGTGTTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGA 1852
Db 358 TTTTGTGTTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGTGA 417
Qy 1853 TGTGTAGTCTCTGTTTTAAATTCAGCAAACTGGAATAATCCATCTAAATTAATGCTT 1912
Db 418 TGTGTAGTCTCTGTTTTAAATTCAGCAAACTGGAATAATCCATCTAAATTAATGCTT 477
Qy 1913 TCTTCCCAAGAGTTTTTAATGATATGCGAGCTTCTCTAATTTGGAGACAAAGCTTA 1972
Db 478 TCTTCCCAAGAGTTTTTAATGATATGCGAGCTTCTCTAATTTGGAGACAAAGCTTA 537
Qy 1973 ATTGACATGATTCATTATATATTTTGTATAGTTACAGTATACGAGTTCAGTATCC 2032
Db 538 ATTGACATGATTCATTATATATTTTGTATAGTTACAGTATACGAGTTCAGTATCC 597
Qy 2033 CTTAGATGAGATGCTTGGGACCAAGAGTGTTCGATTTTTCAGATTTATTTTCGATTTTG 2092
Db 598 CTTAGATGAGATGCTTGGGACCAAGAGTGTTCGATTTTTCAGATTTATTTTCGATTTTG 657
Qy 2093 GAATATTTCCATACATATATGAGAGAGTTGGAAATGGGATTCAGTCTAAATCAATAAA 2152
Db 658 GAATATTTCCATACATATATGAGAGAGTTGGTAAATGGGATTCAGTCTAAATCAATAAA 717
Qy 2153 TTCCTATGTTTGTATATACACCTTATCTGAATAGCCTGAGGTAAATTTTATCAATATT 2212
Db 718 TGCCTTATGTTTGTATATACACCTTATCTGAATAGCCTGAGGTAAATTTTATCAATATT 777
Qy 2213 TTAATATTTTATGCTGAAACAGAGTTTCCGACATTCGACCATCAGAACAGAGT 2272
Db 778 TTAATATTTTATGCTGAAACAGAGTTTCCGACATTCGACCATCAGAACAGAGT 837
Qy 2273 GTCACTATTTCAAGTCAAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGTGATGAGTTTCAT 2332
Db 838 GTCACTATTTCAAGTCAAGTGTCTCAAAAAGTTTCAGATGTTAAGCTGTGATGAGTTTCAT 897
Qy 2333 GCCAGTATCCGAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGAGTTT 2392
Db 898 GCCAGTATCCGAGTACTTTTGGGAAGCCAGACAGAGTGGATCTCTTGGAGCCAGAGTTT 957
Qy 2393 GAGGCCAGACTGCAACACACAGTGCAGACCTCGTTTCT 2429
Db 958 GAGGCCAGACTGCAACACACAGTGCAGACCTCGTTTCT 994

RESULT 10

ADC30760

ID ADC30760 standard; cDNA; 667 BP.

AC ADC30760;

XX 18-DEC-2003 (first entry)

PT Human novel cDNA sequence, SEQ ID NO:842.

DE

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; antihaemic; anticoagulant; thrombolytic; vulnery;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy; chromosome 1; gene; ss.
XX Homo sapiens.
XX WO2003029271-A2.
XX 10-APR-2003.
XX 24-SEP-2002; 2002WO-US030474.
XX 24-SEP-2001; 2001US-0324631P.
XX (HYSE-) HYSEQ INC.
XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX Haley-Vicente D, Drmanac RT;
XX WPI; 2003-371981/35.
XX P-PSDB; ADC31731.

New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer.

Claim 1; SEQ ID NO 842; 1185pp; English.

The invention relates to 971 novel human cDNA sequences (ADC29919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 787 contig sequences corresponding to the cDNA sequences of the invention (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human cDNA sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Query Match 13.5%; Score 506; DB 9; Length 667;

Best Local Similarity 100.0%; Pred. No. 1.4e-165;

Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1875 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 1934
DB 345 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 404

QY 1935 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 1967
DB 405 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 437

RESULT 14
ABV40245
ID ABV40245 standard; cDNA; 437 BP.
XX
AC ABV40245;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 40236.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Pharmacogenomic marker; Gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8129; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
SQ Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;

Query Match 9.1%; Score 342; DB 5; Length 437;
Best Local Similarity 99.7%; Pred. No. 8.6e-109;
Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1575 GTACTGGATATTTTGCATATTAATTCAGTATGCTTCAAAAATTAATAGTTTTTGAC 1634
DB 45 GTACTGGATATTTTGCATATTAATTCAGTATGCTTCAAAAATTAATAGTTTTTGAC 104

QY 1635 ATTGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAATGAAG 1694
DB 105 ATTGCTTTTCTGAGAGAGAAATTTGAAAGTGTCAAAAATAAAAAAGATGAATGAAG 164

QY 1695 CATATATTAATTCCTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAGGATCTGTTC 1754
DB 165 CATATATTAATTCCTCAATTTTTCATTTTCTAGTCAACAGAGAAATCGAGGATCTGTTC 224

QY 1755 AAATATTAGTAAAAATTGAAAAATAAATTTGCTTTATATTTTGTTCACACACTAGTT 1814
DB 225 AAATATTAGTAAAAATTGAAAAATAAATTTGCTTTATATTTTGTTCACACACTAGTT 284

QY 1815 AATTTAACTGTGACTAGTATCTCTACGAGGTGGATGTGTAGTTTCTGGTTTAAAA 1874
DB 285 AATTTAACTGTGACTAGTATCTCTACGAGGTGGATGTGTAGTTTCTGGTTTAAAA 344

QY 1875 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 1934
DB 345 TTCAAGCAAACTGGAAATAATCAATCAATTAATGCTTTCTTCCCAAGAGTTTTTAA 404

QY 1935 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 1967
DB 405 TGATATGCCAGCTTCCTAAATTTGGAGACAAAAG 437

RESULT 15
ABV40577
ID ABV40577 standard; cDNA; 437 BP.
XX
AC ABV40577;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 40568.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW Pharmacogenomic marker; Gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8181; 11750pp; English.

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;

CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 437 BP; 150 A; 63 C; 79 G; 145 T; 0 U; 0 Other;

Query Match 9.1%; Score 342; DB 5; Length 437;
 Best Local Similarity 99.7%; Pred. No. 8.6e-109;
 Matches 392; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1575 GTACTGATATTTTGCATATAAATTCAGTATGCTCAAAAATTAATAGTTTTGAC 1634
 Db |||||
 QY 45 GTACTGATATTTTGCATATAAATTCAGTATGCTCAAAAATTAATAGTTTTGAC 104
 Db |||||
 QY 1635 ATTGGCTTTTCTGAGAGAGAAATGGAAGTGTCAAAAATAAAAAGATGAATGAAG 1694
 Db |||||
 QY 105 ATTGGCTTTTCTGAGAGAGAAATGGAAGTGTCAAAAATAAAAAGATGAATGAAG 164
 Db |||||
 QY 1695 CATATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGATCGAAGATTCGTTC 1754
 Db |||||
 QY 165 CATATATAATTTGCAATTTTTCATTTTCTAGTCAACAGAGATCGAAGATTCGTTC 224
 Db |||||
 QY 1755 AAATATTAGTAAAAATGAAATAAATCTGTGCTTATATTTTGTTCACACACTAGTT 1814
 Db |||||
 QY 225 AAATATTAGTAAAAATGAAATAAATCTGTGCTTATATTTTGTTCACACACTAGTT 284
 Db |||||
 QY 1815 AATTACCTGTGACTAGTTATCTCTACCGAAGTGGATGATGTTCTGGTTTAAAA 1874
 Db |||||
 QY 285 AATTACCTGTGACTAGTTATCTCTACCGAAGTGGATGATGTTCTGGTTTAAAA 344
 Db |||||
 QY 1875 TTCAAGCAACTGGAATAATAATCATCTAATTTATGCTTTCTTCCCAAGAGTTTAA 1934
 Db |||||
 QY 345 TTCAAGCAACTGGAATAATAATCATCTAATTTATGCTTTCTTCCCAAGAGTTTAA 404
 Db |||||
 QY 1935 TGATATGCCAGTTCCTTAATTTGGAGACAAAAG 1967
 Db |||||
 QY 405 TGATATGCCAGTTCCTTAATTTGGAGACAAAAG 437
 Db |||||

RESULT 16
 ABV00935
 ID ABV00935 standard; cDNA; 446 BP.
 XX
 AC ABV00935;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 926.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 PR 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 PR 13-DEC-2000; 2000US-0255281P.
 XX
 PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX
 PI Schlegel R, Endege WO, Monahan JE;
 XX
 DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX
 PS Claim 1; Page 246; 11750pp; English.
 XX
 CC The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 446 BP; 146 A; 68 C; 77 G; 152 T; 0 U; 3 Other;

Query Match 6.5%; Score 244; DB 5; Length 446;
 Best Local Similarity 99.4%; Pred. No. 7e-75;
 Matches 344; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1611 GTTCAAAAATTAATAGTTTTCACATTCGCTTTCTGAGAGAGAAATGAAAGTGTCA 1670
 Db |||||
 QY 80 GTTCAAAAATTAATAGTTTTCACATTCGCTTTCTGAGAGAGAAATGAAAGTGTCA 139
 Db |||||
 QY 1671 AAAATAAAAAAGATGAATGAAGCATATATAATTTGTCAATTTTCAATTTCTAGTCA 1730
 Db |||||
 QY 140 AAAATAAAAAAGATGAATGAAGCATATATAATTTGTCAATTTTCAATTTCTAGTCA 199
 Db |||||
 QY 1731 ACAGAGAAATCGAAGATTCGTTTCAAAATTAATAGTAAAAATTCAAAATAAATCTGTGCTTA 1790
 Db |||||
 QY 200 ACAGAGAAATCGAAGATTCGTTTCAAAATTAATAGTAAAAATTCAAAATAAATCTGTGCTTA 259
 Db |||||
 QY 1791 TATTTTGTTCACACACACTAGTCAATTTAACTGTGACTAGTATCTCTACCGAAGTG 1850
 Db |||||
 QY 260 TATTTTGTTCACACACACTAGTCAATTTAACTGTGACTAGTATCTCTACCGAAGTG 319
 Db |||||
 QY 1851 GATGTGATTTTCTGTTTTTAAATTCAGCAAACTGGAATAATCCATCTAATTTATGC 1910
 Db |||||
 QY 320 GATGTGATTTTCTGTTTTTAAATTCAGCAAACTGGAATAATCCATCTAATTTATGC 379
 Db |||||
 QY 1911 TTCTTTTCCCAAGAGTTTTTTTAAATGATATGCCAGTTCCTTAATTT 1956
 Db |||||
 QY 380 TTCTTTTCCCAAGAGTTTTTTTAAATGATATGCCAGTTCCTTAATTT 425
 Db |||||

RESULT 17
 ABV10104
 ID ABV10104 standard; cDNA; 404 BP.
 XX
 AC ABV10104;
 XX
 DT 13-SEP-2002 (first entry)
 XX
 DE Human prostate expression marker cDNA 10095.
 XX
 KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO200160860-A2.
 XX
 PD 23-AUG-2001.
 XX
 PF 20-FEB-2001; 2001WO-US005171.
 XX
 DR 17-FEB-2000; 2000US-0183319P.

```
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 1617-1618; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 404 BP; 142 A; 54 C; 63 G; 144 T; 0 U; 1 Other;
XX
XX Query Match 5.7%; Score 213; DB 5; Length 404;
XX Best Local Similarity 99.2%; Pred. No. 3.8e-64;
XX Matches 363; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1611 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAAGAGAAATTGAAAGTGTAC 1670
DB 39 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAAGAGAAATTGAAAGTGTAC 98
QY 1671 AAATAAATAAAGATGAATGAGCATATATATTTCAATTTTTCATTTTCTAGTCA 1730
DB 99 AAATAAATAAAGATGAATGAGCATATATATTTCAATTTTTCATTTTCTAGTCA 158
QY 1731 ACAGAGATCGAAGGATCTGTTCAAATATTAGTAAATAATTGAAATAAACTTGTGCTTA 1790
DB 159 ACAGAGATCGAAGGATCTGTTCAAATATTAGTAAATAATTGAAATAAACTTGTGCTTA 218
QY 1791 TATTTTGTTCACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 1850
DB 219 TATTTTGTTCACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 278
QY 1851 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 1910
DB 279 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 338
QY 1911 TTTCTTTCCCAAGAGATTTTTTAATGATATGACCGACTTCTTAATTTGGAGACAAAGCCT 1970
DB 339 TTTCTTTCCCAAGAGATTTTTTAATGATATGACCGACTTCTTAATTTGGAGACAAAGCCT 398
QY 1971 TAAATG 1976
DB 399 TAAATG 404
RESULT 18
ABV31276
ID ABV31276 standard; cDNA; 421 BP.
XX
XX AC ABV31276;
XX DT 16-SEP-2002 (first entry)
```

```
XX Human prostate expression marker cDNA 31267.
XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX Homo sapiens.
XX WO200160860-A2.
XX 23-AUG-2001.
XX 20-FEB-2001; 2001WO-US005171.
XX 17-FEB-2000; 2000US-0183319P.
XX 16-MAR-2000; 2000US-0189862P.
XX 25-MAY-2000; 2000US-0207454P.
XX 09-JUN-2000; 2000US-0211314P.
XX 18-JUL-2000; 2000US-0219007P.
XX 13-DEC-2000; 2000US-0255281P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
XX Claim 1; Page 6740; 11750pp; English.
XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
XX Sequence 421 BP; 143 A; 61 C; 71 G; 145 T; 0 U; 1 Other;
XX
XX Query Match 5.7%; Score 213; DB 5; Length 421;
XX Best Local Similarity 99.2%; Pred. No. 3.8e-64;
XX Matches 363; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
XX
QY 1611 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAAGAGAAATTGAAAGTGTAC 1670
DB 56 GTTCAAAAATTAAATAGTTTTTGACATTGGCTTTTCTGAGAAGAGAAATTGAAAGTGTAC 115
QY 1671 AAATAAATAAAGATGAATGAGCATATATATTTCAATTTTTCATTTTCTAGTCA 1730
DB 116 AAATAAATAAAGATGAATGAGCATATATATTTCAATTTTTCATTTTCTAGTCA 175
QY 1731 ACAGAGATCGAAGGATCTGTTCAAATATTAGTAAATAATTGAAATAAACTTGTGCTTA 1790
DB 176 ACAGAGATCGAAGGATCTGTTCAAATATTAGTAAATAATTGAAATAAACTTGTGCTTA 235
QY 1791 TATTTTGTTCACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 1850
DB 236 TATTTTGTTCACACACACTAGTTAATTTAACTGTGACTAGTTATCTTACCGAAGTG 295
QY 1851 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 1910
DB 296 GATGTGTAGTTTCTGTTTAAATTTCAAGCAAACTGGAATAATCCATCTAATTATGC 355
QY 1911 TTTCTTTCCCAAGAGATTTTTTAATGATATGACCGACTTCTTAATTTGGAGACAAAGCCT 1970
```


QY 1789 TATATTTTGTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGG 1848
Db 258 TATATTTTGTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGG 317
QY 1849 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 1908
Db 318 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 377
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
Db 378 GCTTTCTTTCCCAAGAGTTTTT 400

RESULT 21
ABV39737
ID ABV39737 standard; cDNA; 401 BP.
XX
AC ABV39737;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 39728.
XX
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
KW Homo sapiens.
OS
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 8048; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 401 BP; 136 A; 58 C; 73 G; 134 T; 0 U; 0 Other;

Query Match 5.4%; Score 203; DB 5; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.1e-60; Indels 0; Gaps 0;
Matches 203; Conservative 0; Mismatches 0;

QY 1729 CAACAGAGATCGAAGGATTCGTTCAAATATTAGTAAATTTGAAATAAATCTGTGCT 1788
Db 199 CAACAGAGATCGAAGGATTCGTTCAAATATTAGTAAATTTGAAATAAATCTGTGCT 258
QY 1789 TATATTTTGTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGG 1848
Db 259 TATATTTTGTGCAACACACTAGTAAATTTAACTGTGACTAGTATCTCTACCGAAGG 318
QY 1849 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 1908
Db 319 TGGATGTGAGTCTTCTGTTTAAATTCAGCAAACTGGAAATATCCATCTAATTAT 378
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
Db 379 GCTTTCTTTCCCAAGAGTTTTT 401

RESULT 22
ABV30769
ID ABV30769 standard; cDNA; 401 BP.
XX
AC ABV30769;
XX
DT 16-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 30760.
XX
DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
KW Homo sapiens.
OS
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 6663-6664; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 401 BP; 136 A; 58 C; 73 G; 134 T; 0 U; 0 Other;

Query Match 5.4%; Score 203; DB 5; Length 401;

```
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1729 CAACAGAGATCGAAGGATTCCTGTCATATATTAGTAAATAATTTGAAAATAAACTTGTGCT 1788
DB 199 CAACAGAGATCGAAGGATTCCTGTCATATATTAGTAAATAATTTGAAAATAAACTTGTGCT 258
QY 1789 TATATTTGTTTCCAAACACACTAGTTAAATTAACCTGTGACTAGTTATCTTACCGAAGG 1848
DB 259 TATATTTGTTTCCAAACACACTAGTTAAATTAACCTGTGACTAGTTATCTTACCGAAGG 318
QY 1849 TGGATGTGATGTTCTGTTTAAATTCAGCAAACTGGAATAATCAATCTAATAT 1908
DB 319 TGGATGTGATGTTCTGTTTAAATTCAGCAAACTGGAATAATCAATCTAATAT 378
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
DB 379 GCTTTCTTTCCCAAGAGTTTTT 401

RESULT 23
ABV39958
ID ABV39958 standard; cDNA; 401 BP.
XX AC ABV39958;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 39949.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 8083; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient;
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 401 BP; 136 A; 58 C; 73 G; 134 T; 0 U; 0 Other;
Query Match 5.4%; Score 203; DB 5; Length 401;
Best Local Similarity 100.0%; Pred. No. 1.1e-60;
Matches 203; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1729 CAACAGAGATCGAAGGATTCCTGTCATATATTAGTAAATAATTTGAAAATAAACTTGTGCT 1788
DB 199 CAACAGAGATCGAAGGATTCCTGTCATATATTAGTAAATAATTTGAAAATAAACTTGTGCT 258
QY 1789 TATATTTGTTTCCAAACACACTAGTTAAATTAACCTGTGACTAGTTATCTTACCGAAGG 1848
DB 259 TATATTTGTTTCCAAACACACTAGTTAAATTAACCTGTGACTAGTTATCTTACCGAAGG 318
QY 1849 TGGATGTGATGTTCTGTTTAAATTCAGCAAACTGGAATAATCAATCTAATAT 1908
DB 319 TGGATGTGATGTTCTGTTTAAATTCAGCAAACTGGAATAATCAATCTAATAT 378
QY 1909 GCTTTCTTTCCCAAGAGTTTTT 1931
DB 379 GCTTTCTTTCCCAAGAGTTTTT 401

RESULT 24
ABV30990
ID ABV30990 standard; cDNA; 296 BP.
XX AC ABV30990;
XX DT 16-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 30981.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX DR WPI; 2001-662795/76.
XX PT Novel isolated nucleic acid molecule associated with cancerous state of
XX PT prostate cells and correlating with presence of prostate cancer, useful
XX PT for detecting presence of prostate cancer, stage of prostate cancer.
XX PS Claim 1; Page 6697; 11750pp; English.
XX CC The invention relates to an isolated nucleic acid molecule (I) comprising
XX CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
XX CC specification or its complement. (I) is useful for: (a) assessing whether
XX CC a patient is afflicted with prostate cancer; (b) monitoring the
XX CC progression of prostate cancer in a patient; (c) assessing the efficacy
XX CC of a test compound to inhibit prostate cancer in a patient;
XX CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
XX CC (e) assessing the prostate cell carcinogenic potential of a compound; (g)
XX CC determining whether prostate cancer has metastasized in a patient; (h)
XX CC assessing the aggressiveness or indolence of prostate cancer in a patient
XX CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
```


Db 193 TAAAAATTGAAATAAATCTGTCTATATTTGTTTGCACACACTAGTATTAATTAACC 252
QY 1824 TG 1825
Db 253 TG 254

RESULT 27

ABV56428
ID ABV56428 standard; cDNA; 381 BP.
XX AC ABV56428;
XX DT 17-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 56419.
XX DE Human prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
XX KW pharmacogenomic marker; gene; ss.
XX OS Homo sapiens.
XX PN WO200160860-A2.
XX PD 23-AUG-2001.
XX PF 20-FEB-2001; 2001WO-US005171.
XX PR 17-FEB-2000; 2000US-0183319P.
XX PR 16-MAR-2000; 2000US-0189862P.
XX PR 25-MAY-2000; 2000US-0207454P.
XX PR 09-JUN-2000; 2000US-0211314P.
XX PR 18-JUL-2000; 2000US-0219007P.
XX PR 13-DEC-2000; 2000US-0255281P.
XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX PI Schlegel R, Endege WO, Monahan JE;
XX WPI; 2001-662795/76.
XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX Claim 1; Page 10887; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 381 BP; 165 A; 67 C; 70 G; 78 T; 0 U; 1 Other;

Query Match 1.4%; Score 52; DB 5; Length 381;
Best Local Similarity 100.0%; Pred. No. 2e-08;
Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3711 TATAAATGAAAAAATAAATCTGTCTATATTTGTTTGCACACACTAGTATTAATTAACC 3762
Db 230 TATAAATGAAAAAATAAATCTGTCTATATTTGTTTGCACACACTAGTATTAATTAACC 281

RESULT 28

Human prostate expression marker cDNA 57384.
Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

AAI90783
ID AAI90783 standard; cDNA; 345 BP.
XX AC AAI90783;
XX DT 06-NOV-2001 (first entry)
XX DE Human polynucleotide SEQ ID NO 10843.
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation; ss.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX WPI; 2001-514838/56.
XX DR P-PSDB; AAO10852.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 1; SEQ ID NO 10843; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 345 BP; 142 A; 54 C; 69 G; 79 T; 0 U; 1 Other;

Query Match 1.3%; Score 50; DB 4; Length 345;
Best Local Similarity 100.0%; Pred. No. 1e-07;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAAATAAATCTGTCTATATTTGTTTGCACACACTAGTATTAATTAACC 3762
Db 278 TAAATGAAAAAATAAATCTGTCTATATTTGTTTGCACACACTAGTATTAATTAACC 327

RESULT 29

ABV57393
ID ABV57393 standard; cDNA; 424 BP.
XX AC ABV57393;
XX DT 17-SEP-2002 (first entry)
XX DE Human prostate expression marker cDNA 57384.
XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

KW pharmacogenomic marker; gene; ss.

OS Homo sapiens.

XX WO200160860-A2.

XX PD 23-AUG-2001.

XX PF 20-FEB-2001; 2001WO-US005171.

XX PR 17-FEB-2000; 2000US-0183319P.

PR 16-MAR-2000; 2000US-0189862P.

PR 25-MAY-2000; 2000US-0207454P.

PR 09-JUN-2000; 2000US-0211314P.

PR 18-JUL-2000; 2000US-0219007P.

PR 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX PI Schlegel R, Endege WO, Monahan JB;

XX DR WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 11041; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 424 BP; 100 A; 114 C; 73 G; 136 T; 0 U; 1 Other;

Query Match 1.3%; Score 50; DB 5; Length 424;

Best Local Similarity 100.0%; Pred. No. 9.7e-08;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762

Db 47 TAAATGAAA 96

RESULT 30

AAI83050

ID AAI83050 standard; cDNA; 449 BP.

XX AC AAI83050;

XX 06-NOV-2001 (first entry)

XX Human polynucleotide SEQ ID NO 3110.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy; vaccine; peptide therapy; stem cell growth factor; haematopoiesis; tissue growth factor; immunomodulatory; cancer; leukaemia; nervous system disorders; arthritis; inflammation; ss.

OS Homo sapiens.

XX WO200164835-A2.

XX PD 07-SEP-2001.

XX

PF

XX

PR

PR

XX

PA

XX

PI

XX

DR

DR

XX

PT

PT

XX

PS

XX

XX

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

CC

26-FEB-2001; 2001WO-US004927.

28-FEB-2000; 2000US-00515126.

18-MAY-2000; 2000US-00577409.

(HYSE-) HYSEQ INC.

Tang YT, Liu C, Drmanac RT;

WPI; 2001-514838/56.

P-PSDB; AAO03119.

Isolated nucleic acids and polypeptides, useful for preventing diagnosing and treating e.g. leukemia, inflammation and immune disorders.

Claim 1; SEQ ID NO 3110; 1399pp + Sequence Listing; English.

The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to cytokine, cell proliferation or cell differentiation or which may induce production of other cytokines in other cell populations. The polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, haematopoiesis regulating activity, tissue growth factor activity, immunomodulatory activity and activin/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 449 BP; 184 A; 55 C; 100 G; 109 T; 0 U; 1 Other;

Query Match 1.3%; Score 50; DB 4; Length 449;

Best Local Similarity 100.0%; Pred. No. 9.5e-08;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762

Db 125 TAAATGAAA 174

RESULT 31

AAFI5960

ID AAFI5960 standard; cDNA; 1702 BP.

XX AC AAFI5960;

XX 13-MAR-2001 (first entry)

XX Human prostate cancer antigen nucleotide sequence SEQ ID NO:395.

Human; prostate cancer; prostate cancer antigen; detection; diagnosis; neuroprotective; cytostatic; cardioprotective; immunomodulatory; muscular; vulvular; gastrointestinal; nephrotropic; anti-infective; gynaecological; antibacterial; gene therapy; neural; immune; reproductive; renal; gastrointestinal; pulmonary; cardiovascular; proliferative disorder; wound; infectious disease; ss.

OS Homo sapiens.

XX WO200055174-A1.

XX PD 21-SEP-2000.

XX 08-MAR-2000; 2000WO-US005988.

XX 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX (ROSE/) ROSEN C A.

XX


```
PI Rosen CA, Ruben SM;
XX
DR WPI: 2000-587513/55.
DR P-PSDB; AAB56757.
XX
XX Prostate cancer associated gene sequences, referred to as prostate cancer
PT antigens, useful for treatment, prevention, and diagnosis of disorders
PT such as prostate cancer.
XX
PS Claim 1; Page 885; 2338pp; English.
XX
CC AAF1566 to AAF16505 encode the human prostate cancer associated
CC proteins, called prostate cancer antigens, given in AAB56363 to AAB57302.
CC The prostate cancer antigens can have neuroprotective, cytostatic,
CC cardioactive, immunomodulatory, muscular, vulnery, gastrointestinal,
CC nephrotropic, antiinfective, gynaecological and antibacterial activities,
CC and can be used in gene therapy. The prostate cancer antigen
CC polynucleotides may be used for detection of prostate cancer, chromosome
CC identification, as chromosome markers, and for numerous other diagnostic
CC or research purposes. The prostate cancer antigens may be used to treat
CC disorders such as neural, immune, muscular, reproductive,
CC gastrointestinal, pulmonary, cardiovascular, renal, and proliferative
CC disorders, wounds, and infectious diseases. AAF16506 to AAF16514 to
CC AAB57303 represent sequences used in the exemplification of the present
CC invention
XX
SQ Sequence 1702 BP; 608 A; 284 C; 302 G; 500 T; 0 U; 8 Other;
Query Match 1.3%; Score 50; DB 3; Length 1702;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1588 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1637
RESULT 32
ADD18824
ID ADD18824 standard; DNA; 1820 BP.
XX
AC ADD18824;
XX
DT 15-JAN-2004 (first entry)
XX
DE Human disease related protein DNA sequence SeqID256.
XX
KW human; disease state; cytostatic; antiinflammatory; ophthalmological;
KW antiarteriosclerotic; vulnery; gene therapy;
KW hypoxia-regulated condition; tumorigenesis; angiogenesis; apoptosis;
KW inflammation; erythropoiesis; glycolysis; gluconeogenesis;
KW glucose transportation; catecholamine synthesis; iron transport;
KW nitric oxide synthesis; cancer; ischaemic condition; reperfusion injury;
KW retinopathy; neonatal stress; pre-eclampsia; atherosclerosis;
KW inflammatory condition; wound healing; gene; ds.
XX
OS Homo sapiens.
XX
PN WO2003018621-A2.
XX
PD 06-MAR-2003.
XX
PF 23-AUG-2002; 2002WO-GB003892.
XX
PR 23-AUG-2001; 2001GB-00020558.
XX
PR 05-OCT-2001; 2001GB-00024037.
XX
PA (OXFO-) OXFORD BIOMEDICA UK LTD.
XX
PI Kingsman SM, White J, Ward NR, Harris RA, Naylor S, Mundy CR;
DR WPI: 2003-290046/28.
DR P-PSDB; ADD18823.
XX
XX New substantially purified polypeptide, useful for diagnosing or treating
PT a hypoxia-regulated condition, such as cancer, ischemia, reperfusion
PT injury, retinopathy, pre-eclampsia, atherosclerosis, inflammation, or
PT wound healing.
XX
XX Claim 27; SEQ ID NO 256; 424pp; English.
XX
XX This invention relates to novel human genes and gene product which are
CC implicated in certain disease states. Compounds which modulate the
CC proteins of the invention may have cytostatic, antiinflammatory,
CC ophthalmological, antiarteriosclerotic or vulnery activities. The
CC sequences of the invention may be useful for gene therapy. The invention
CC may be useful for diagnosing or treating a hypoxia-regulated condition,
CC such as tumorigenesis, angiogenesis, apoptosis, inflammation,
CC erythropoiesis, or the biological response to hypoxia conditions
CC including processes such as glycolysis, gluconeogenesis, glucose
CC transportation, catecholamine synthesis, iron transport or nitric oxide
CC synthesis. The disease includes cancer, ischaemic conditions, reperfusion
CC injury, retinopathy, neonatal stress, pre-eclampsia, atherosclerosis,
CC inflammatory conditions or wound healing. The present sequence is that of
CC a disease related protein encoding DNA sequence of the invention.
XX
SQ Sequence 1820 BP; 451 A; 485 C; 472 G; 412 T; 0 U; 0 Other;
Query Match 1.3%; Score 50; DB 9; Length 1820;
Best Local Similarity 100.0%; Pred. No. 7.3e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 1719 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1768
RESULT 33
ACF12861
ID ACF12861 standard; cDNA; 3602 BP.
XX
AC ACF12861;
XX
DT 10-SEP-2003 (first entry)
XX
DE Human cervical cancer cell marker encoding cDNA SEQ ID NO:67.
XX
KW Human; cervical cancer; cervical cancer marker; cancer therapy;
KW detection; gene therapy; vaccine; gene; ss.
XX
OS Homo sapiens.
XX
PN WO2002101075-A2.
XX
PD 19-DEC-2002.
XX
PF 12-JUN-2002; 2002WO-US018638.
XX
PR 13-JUN-2001; 2001US-0298155P.
XX
PR 13-JUN-2001; 2001US-0298159P.
XX
PR 14-NOV-2001; 2001US-0335936P.
XX
PA (MILL-) MILLENNIUM PHARM INC.
XX
PI Schlegel R, Chen Y, Zhao X, Monahan JE, Kamatkar S;
PI Gannavarapu M, Glatt K, Hoersch S;
XX
XX WPI: 2003-156967/15.
XX
XX New isolated nucleic acid molecule useful for detecting, characterizing,
PT preventing and treating human cervical cancers, in various prognostic and
PT diagnostic assays, in pharmacogenomics and in monitoring clinical trials.
XX
XX Claim 1; Page 218-219; 386pp; English.
XX
XX ACF12828 to ACF12947 encode the human cervical cancer marker proteins (I)
```

CC given in ABR92047 to ABR92164. A higher level of expression of (I) than
CC normal indicates the presence of cervical cancer. Also described: (1) a
CC vector (II) containing (I); (2) a host cell (III) containing (I); and (3)
CC assessing (MI) whether a patient is afflicted with cervical cancer,
CC comprising comparing the level of expression of a marker in a patient's
CC sample, and the normal level of expression of the marker in a control non
CC -cervical cancer sample, where a significant increase in the level of
CC expression of the marker in the patient's sample relative to that in the
CC control sample is an indication that the patient is afflicted with
CC cervical cancer. (I) has cytostatic activity, and can be used in gene
CC therapy and in vaccines. (I) is useful in detecting, characterising,
CC preventing and treating human cervical cancers. (I) may also be used in
CC various prognostic and diagnostic assays, pharmacogenomics and in
CC monitoring clinical trials

XX Sequence 3602 BP; 790 A; 1062 C; 1017 G; 730 T; 0 U; 3 Other;

Query Match 1.3%; Score 50; DB 7; Length 3602;
Best Local Similarity 100.0%; Pred. No. 6.3e-08;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
DB 3501 TAAATGAAAAA 3550

RESULT 34
AA87412
ID AA87412 standard; cDNA; 3607 BP.

AC AA87412;

DT 08-OCT-1999 (first entry)

DE Hepatocellular carcinoma marker gene L4 cDNA.

XX Hepatocellular carcinoma; tumour; cancer; diagnosis; marker; probe;
KW human; hepatitis B virus; x antigen; HBxAg; L4 gene; ss.

XX Homo sapiens.

XX WO9939200-A1.

XX 05-AUG-1999.

XX 27-JAN-1999; 99WO-US001894.

XX 29-JAN-1998; 98US-0072938P.

XX (UYJE-) UNIV JEFFERSON THOMAS.

XX Feitelson MA;

XX WPI; 1999-469371/39.

XX New method for diagnosing hepatocellular carcinoma.

XX Claim 3; Page 43-44; 52pp; English.

XX This sequence represents a gene, designated L4, whose expression is
CC activated in human hepatoblastoma HepG2 cells in the presence of
CC hepatitis B virus x antigen (HBxAg) compared to expression in HepG2 cells
CC in the absence of HBxAg. The full-length cDNA was obtained from HepG2
CC cells using 5' and 3' RACE. No sequence homology was found between L4 and
CC any previously known sequence. Synthetic peptides that represent probable
CC antigenic determinants of the L4 protein are provided in AA06538 and
CC AA06539. L4 is 1 of 10 genes (see AA87402-11) that were initially
CC identified by PCR select cDNA subtraction as showing differential
CC expression in HBxAg(+) and HBxAg(-) cells. It can be used as a molecular
CC marker for hepatocellular carcinoma (HCC). In a claimed method for
CC detecting HCC, a liver tissue sample is obtained from a patient, and the
CC level of expression of 1 or more marker genes, such as L4, in the sample
CC is assessed. A reduction in the level of expression of the marker genes

CC in the sample as compared to the expression level in noncancerous liver
CC tissue is indicative of HCC. The method allows screening for the risk of
CC disease and early diagnosis before tumours develop

SQ Sequence 3607 BP; 875 A; 938 C; 1058 G; 736 T; 0 U; 0 Other;

Query Match 1.3%; Score 50; DB 2; Length 3607;

Best Local Similarity 100.0%; Pred. No. 6.3e-08;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

DB 3531 TAAATGAAAAA 3580

RESULT 35

ABV48235

ID ABV48235 standard; cDNA; 265 BP.

AC ABV48235;

DT 17-SEP-2002 (first entry)

DE Human prostate expression marker cDNA 48226.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-652795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

PS Claim 1; Page 9463; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV0010-ABV6213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

SQ Sequence 265 BP; 159 A; 27 C; 31 G; 48 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 5; Length 265;

Best Local Similarity 100.0%; Pred. No. 2.3e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 3714 AAATGAAAAA... 3762
Db 100 AAATGAAAAA... 148

RESULT 36
ABV18451
ID ABV18451 standard; cDNA; 287 BP.
XX
AC ABV18451;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 18442.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
XX
PR 16-MAR-2000; 2000US-0189862P.
XX
PR 25-MAY-2000; 2000US-0207454P.
XX
PR 09-JUN-2000; 2000US-0211314P.
XX
PR 18-JUL-2000; 2000US-0219007P.
XX
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3040-3041; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 287 BP; 147 A; 29 C; 29 G; 50 T; 0 U; 32 Other;

Query Match 1.3%; Score 49; DB 5; Length 287;
Best Local Similarity 100.0%; Pred. No. 2.3e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 62 AAATGAAAAA... 110

RESULT 37
AAI84445
ID AAI84445 standard; cDNA; 381 BP.
```

```
XX AAI84445;
XX 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 4505.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
PN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
PF 26-FEB-2001; 2001WO-US004927.
XX
PR 28-FEB-2000; 2000US-00515126.
XX
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
XX
DR P-PSDB; AAO04514.
XX
PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX
PS Claim 1; SEQ ID NO 4505; 1399pp + Sequence Listing; English.
XX
CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 381 BP; 154 A; 66 C; 81 G; 80 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 4; Length 381;
Best Local Similarity 100.0%; Pred. No. 2.2e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
Db 46 AAATGAAAAA... 94

RESULT 38
ABZ78073/C
ID ABZ78073 standard; cDNA; 476 BP.
XX
AC ABZ78073;
XX
DT 22-APR-2003 (first entry)
XX
DE Human breast specific nucleic acid #87.
XX
KW Human; breast specific nucleic acid; BSN; breast; cytostatic;
KW gene therapy; vaccines; lung cancer; breast cancer;
KW breast specific polypeptide; BSP; gene; ss.
```

XX Homo sapiens.
 XX WO200268645-A2.
 XX PD 06-SEP-2002.
 XX PF 20-NOV-2001; 2001WO-US045151.
 XX PR 20-NOV-2000; 2000US-0249992P.
 XX PA (DIAD-) DIADEXUS INC.
 XX PI Salceda S, Macina RA, Recipon H, Cafferkey R, Sun Y, Liu C;
 XX PI Turner LR;
 XX WPI; 2002-713379/77.
 XX New breast specific genes and proteins, useful in gene therapy or as
 PT vaccines for treating breast cancer or non-cancerous breast diseases, as
 PT well as for diagnosing, monitoring or staging these diseases.
 XX Claim 1; Page 198-199; 277pp; English.
 XX The invention relates to a novel isolated breast specific nucleic acid
 CC molecule. The polypeptides of the invention have cytostatic activity. The
 CC novel nucleic acids and polypeptides may have a use in gene therapy, and
 CC as vaccines. The breast specific nucleic acid and polypeptide are useful
 CC for diagnosing and monitoring the presence and metastases of lung cancer
 CC in a patient. The antibody that specifically binds to the breast specific
 CC polypeptide is useful for determining the presence of a breast specific
 CC protein in a sample, as well as for treating a patient with breast
 CC cancer, particularly by inducing an immune response against the breast
 CC cancer cell expressing the breast specific nucleic acid molecule or
 CC polypeptide. In particular, these breast specific genes and proteins are
 CC useful for identifying, diagnosing, monitoring, staging, imaging and
 CC treating breast cancer and non-cancerous disease states in breast tissue.
 CC These are also useful in gene therapy, production of transgenic animals
 CC and cells, and in the production of engineered breast tissue for
 CC treatment and research. The sequences shown in AB277987-AB278101
 CC represent the novel human breast specific nucleic acid molecules of the
 CC invention
 XX Sequence 476 BP; 116 A; 88 C; 90 G; 182 T; 0 U; 0 Other;
 SQ
 Query Match 1.3%; Score 49; DB 6; Length 476;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 DB 126 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 78
 RESULT 39
 ABV58273
 ID ABV58273 standard; cDNA; 497 BP.
 XX AC ABV58273;
 XX DT 13-SEP-2002 (first entry)
 XX Human prostate expression marker cDNA 58264.
 DE Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 KW pharmacogenomic marker; gene; ss.
 XX Homo sapiens.
 XX WO200160860-A2.
 XX PD 23-AUG-2001.

PF 20-FEB-2001; 2001WO-US005171.
 XX 17-FEB-2000; 2000US-0183319P.
 PR 16-MAR-2000; 2000US-0189862P.
 PR 25-MAY-2000; 2000US-0207454P.
 PR 09-JUN-2000; 2000US-0211314P.
 PR 18-JUL-2000; 2000US-0219007P.
 XX 13-DEC-2000; 2000US-0255281P.
 XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX Schlegel R, Endege WO, Monahan JB;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 11189; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
 XX
 SQ Sequence 497 BP; 235 A; 77 C; 69 G; 115 T; 0 U; 1 Other;
 Query Match 1.3%; Score 49; DB 5; Length 497;
 Best Local Similarity 100.0%; Pred. No. 2.1e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 OY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 DB 252 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 300
 RESULT 40
 AAT72173
 ID AAT72173 standard; cDNA to mRNA; 882 BP.
 XX AC AAT72173;
 XX DT 25-FEB-1998 (first entry)
 XX Alzheimer's disease DNA sequence from plasmid pGCS1180.
 DE Human; brain; Alzheimer's disease; diagnosis; antibody; expression; ds.
 KW Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 97..198
 FT /*tag= a
 FT /note= "protein encoded by DNA from the human brain which
 FT has different expression from the normal protein when in
 FT an Alzheimer's disease patient"
 XX WO9721807-A1.
 XX 19-JUN-1997.
 XX 12-DEC-1996; 96WO-JP003630.
 XX 12-DEC-1995; 95JP-00322745.

XX	CC	The polynucleotide sequences given in AAC79681 to AAC79730 encode the
CC	CC	human secreted proteins given in AAB44596 to AAB44645. AAB44646 to
CC	CC	AAB44693 represent human secreted polypeptide sequences and proteins
CC	CC	homologous to them, which are given in the exemplification of the present
CC	CC	invention. Human secreted proteins have activities based on the tissues
CC	CC	and cells the genes are expressed in. Examples of activities include:
CC	CC	cytostatic; immunosuppressive; neurotropic; neuroprotective; antiviral;
CC	CC	antiallergic; hepatotropic; antidiabetic; antifungal; antiparasitic; and
CC	CC	vulnerable; anticonvulsant; antibacterial; antifungal; antiparasitic; and
CC	CC	cardiant. The polynucleotides and polypeptides are useful for preventing
CC	CC	treating or ameliorating a medical condition in e.g. humans, mice,
CC	CC	rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
CC	CC	can also be used as a food additive or preservative to increase or
CC	CC	decrease storage capabilities. The polynucleotides are useful for
CC	CC	chromosome identification. They are also useful as probes for diagnosing
CC	CC	a disorder related to the female reproductive system, particularly breast
CC	CC	and/or ovary cancer. They are also useful in the gene therapy of breast
CC	CC	and ovarian cancer. The nucleic acids, protein, antibodies, agonists and
CC	CC	antagonists from the present invention are useful in the diagnosis,
CC	CC	treatment and prevention of: cancer; immune disorders; cardiovascular
CC	CC	disorders; wound healing; neurological diseases; and infectious diseases
CC	CC	AAC79672 to AAC79680 and AAB44595 represent sequences used in the
CC	CC	exemplification of the present invention
XX	XX	
XX	XX	Sequence 903 BP; 350 A; 162 C; 173 G; 218 T; 0 U; 0 Other;
XX	XX	
XX	XX	Query Match 1.3%; Score 49; DB 3; Length 903;
XX	XX	Best Local Similarity 100.0%; Fred.No. 1.0e-07;
XX	XX	Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps
XX	XX	
QY	3714	AAATGAAA 3762
DB	788	AAATGAAA 836
XX	XX	
XX	XX	RESULT 42
XX	XX	AAH33675
ID	AAH33675	standard; cDNA; 1057 BP.
XX	XX	
XX	XX	AAH33675;
XX	XX	
XX	XX	03-SEP-2001 (first entry)
XX	XX	
XX	XX	Human colon cancer antigen encoding cDNA SEQ ID NO:731.
XX	XX	
XX	XX	Human; colon cancer; colon cancer antigen; diagnosis; detection;
KW	XX	colorectal carcinoma; ss.
XX	XX	
XX	XX	Homo sapiens.
XX	XX	
XX	XX	W0200122920-A2.
XX	XX	
XX	XX	05-APR-2001.
XX	XX	
XX	XX	28-SEP-2000; 2000WO-US026524.
XX	XX	
XX	XX	29-SEP-1999; 99US-0157137P.
PR	XX	03-NOV-1999; 99US-0163280P.
XX	XX	
XX	XX	(HUMA-) HUMAN GENOME SCI INC.
XX	XX	
XX	XX	Ruben SM, Barash SC, Birse CE, Rosen CA;
XX	XX	
XX	XX	WPI; 2001-235357/24.
DR	XX	P-PSDB; AAG74244.
XX	XX	
XX	XX	Nucleic acids encoding 4277 human colon cancer-associated polypeptides,
PT	XX	useful for preventing, diagnosing and/or treating colorectal cancers.
XX	XX	
XX	XX	Claim 1; Page 2724; 9803pp; English.
XX	XX	
XX	XX	AAH32943 to AAH37195 and AAG73514 to AAG77788 represent human colon

CC cancer-associated nucleic acid molecules (N) and proteins (P), where the
 CC proteins are collectively known as colon cancer antigens. The colon
 CC cancer antigens have cytostatic activity and can be used in gene therapy
 CC and vaccine production. N and P may be used in the prevention, diagnosis
 CC and treatment of diseases associated with inappropriate P expression. For
 CC example, N and P may be used to treat disorders associated with decreased
 CC expression by rectifying mutations or deletions in a patient's genome
 CC that affect the activity of P by expressing inactive proteins or to
 CC supplement the patient's own production of P. Additionally, N may be used
 CC to produce the colon cancer-associated Ps, by inserting the nucleic acids
 CC into a host cell and culturing the cell to express the proteins. N and P
 CC can be used in the prevention, diagnosis and treatment of colorectal
 CC carcinomas and cancers. AAH37196 to AAH37204 and AAB77789 represent
 CC sequences used in the exemplification of the present invention. N.B.
 CC Pages 666 to 682 and page 7053 of the sequence listing were missing at
 CC time of publication, meaning no sequences are present for SEQ ID NO:1027
 CC to 1052, 7921 and 7922

XX
 SQ Sequence 1057 BP; 318 A; 210 C; 207 G; 320 T; 0 U; 2 Other;

Query Match 1.3%; Score 49; DB 4; Length 1057;
 Best Local Similarity 100.0%; Pred. No. 1.8e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
 DB 1006 AAATGAAAAA... 1054

RESULT 43
 AAC79719
 ID AAC79719 standard; cDNA; 1263 BP.

AC AAC79719;

DT 12-FEB-2001 (first entry)

DE Human secreted protein gene 39 SEQ ID NO:49.

KW Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
 KW neutropic; neuroprotective; antiviral; antiallergic; hepatotropic;
 KW antidiabetic; antiinflammatory; antitumor; anticonvulsant;
 KW antibacterial; antifungal; antiparasitic; cardiant; gene therapy;
 KW food additive; preservative; chromosome identification; cancer;
 KW immune disorder; cardiovascular disorder; neurological disease;
 KW wound healing; infectious disease; ss.

XX Homo sapiens.

XX WO200058339-A2.

XX PD 05-OCT-2000.

XX PF 22-MAR-2000; 2000WO-US007440.

XX PR 26-MAR-1999; 99US-0126503P.

XX PR 17-DEC-1999; 99US-0172409P.

XX PA (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM, Komatsoulis G;

XX DR WPI; 2000-594637/56.

XX DR P-PSDB; AAB44634.

XX Fifty nucleic acid molecules encoding human secreted proteins, useful in
 PT the prevention, treatment and diagnosis of cancer, immune disorders,
 PT cardiovascular disorders and neurological diseases.

XX Claim 1; Page 360; 410pp; English.

XX The polynucleotide sequences given in AAC79681 to AAC79730 encode the
 CC human secreted proteins given in AAB44596 to AAB44645. AAB44646 to

CC AAB44693 represent human secreted polypeptide sequences and proteins
 CC homologous to them, which are given in the exemplification of the present
 CC invention. Human secreted proteins have activities based on the tissues
 CC and cells the genes are expressed in. Examples of activities include:
 CC cytostatic; immunosuppressive; neutropic; neuroprotective; antiviral;
 CC antiallergic; hepatotropic; antidiabetic; antiinflammatory; antitumor;
 CC vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic; and
 CC cardiant. The polynucleotides and polypeptides are useful for preventing,
 CC treating or ameliorating a medical condition in e.g. humans, mice,
 CC rabbits, goats, horses, cats, dogs, chickens or sheep. The polypeptides
 CC can also be used as a food additive or preservative to increase or
 CC decrease storage capabilities. The polynucleotides are useful for
 CC chromosome identification. They are also useful as probes for diagnosing
 CC a disorder related to the female reproductive system, particularly breast
 CC and/or ovary cancer. They are also useful in the gene therapy of breast
 CC and ovarian cancer. The nucleic acids, protein, antibodies, agonists and
 CC antagonists from the present invention are useful in the diagnosis,
 CC treatment and prevention of: cancer; immune disorders; cardiovascular
 CC disorders; wound healing; neurological diseases; and infectious diseases.
 CC AAC79672 to AAC79680 and AAB44595 represent sequences used in the
 XX exemplification of the present invention

SQ Sequence 1263 BP; 306 A; 383 C; 323 G; 245 T; 0 U; 6 Other;

Query Match 1.3%; Score 49; DB 3; Length 1263;
 Best Local Similarity 100.0%; Pred. No. 1.7e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA... 3762
 DB 1189 AAATGAAAAA... 1237

RESULT 44

AAC77616

ID AAC77616 standard; cDNA; 1373 BP.

XX AC AAC77616;

XX DT 08-FEB-2001 (first entry)

XX Human cancer associated gene sequence SEQ ID NO:10.

KW Human; cancer associated gene; cancer antigen; detection; cancer;
 KW diagnosis; cytostatic; proliferative; vulnary; immunomodulator;
 KW antidiabetic; antiasthmatic; antirheumatic; antiarthritic; antiviral;
 KW antiinflammatory; antithyroid; antiallergic; antibacterial; cardiant;
 KW dermatological; neuroprotective; thrombolytic; coagulant; neutropic;
 KW vasotropic; antipsoriatic; antiangiogenic; gene therapy; inflammation;
 KW immune disorder; haematopoietic cell disorder; autoimmune disorder;
 KW allergic reaction; graft versus host disease; organ rejection;
 KW haemostatic; thrombolytic; cardiovascular disorder; infection;
 KW neurological disease; drug screening; ss.

XX Homo sapiens.

XX WO200055350-A1.

XX PD 21-SEP-2000.

XX PF 08-MAR-2000; 2000WO-US005882.

XX PR 12-MAR-1999; 99US-0124270P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX PI Rosen CA, Ruben SM;

XX WPI; 2000-587533/55.

XX DR P-PSDB; AAB43407.

XX Novel isolated nucleic acids comprising sequences encoding peptides
 PT useful for treating or diagnosing e.g. cancer.

XX Claim 1; Page 616-617; 2352pp; English.

XX AAC77607 to AAC78448 encode the human cancer associated proteins given in

XX ABA43398 to ABA44239. The proteins can have activities based on the

XX tissues and cells the genes are expressed in. Example of activities

XX include: cytostatic; proliferative; vulnerary; immunomodulator;

XX antidiabetic; antiaesthetic; antirheumatic; antithrombotic;

XX antiinflammatory; antithyroid; antiallergic; antibacterial; antiviral;

XX dermatological; neuroprotective; cardiant; thrombolytic; coagulant;

XX neotropic; vasotropic; antipsoriatic and antiangiogenic. The

XX polynucleotides and polypeptides can be used for preventing, treating or

XX ameliorating medical conditions and diagnosing pathological conditions.

XX Polynucleotides, polypeptides, antibodies, agonists and antagonists from

XX the present invention may be used to treat immune disorders by activating

XX or inhibiting the proliferation, differentiation or mobilisation of

XX immune cells, to treat disorders of haematopoietic cells, autoimmune

XX disorders, allergic reactions, graft versus host disease and organ

XX rejection, modulate haemostatic or thrombolytic activity, modulate

XX inflammation, cancers, cardiovascular disorders, neurological disease and

XX bacterial or viral infections. The peptides, nucleotides, antibodies,

XX agonists and antagonists may be also be used in drug screens. AAC78449 to

XX AAC78457 and ABA44240 represent sequences used in the exemplification of

XX the present invention

SQ Sequence 1373 BP; 398 A; 362 C; 370 G; 241 T; 0 U; 2 Other;

Query Match 1.3%; Score 49; DB 3; Length 1373;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

DB 1294 AAATGAAAAA 1342

RESULT 45

ABV94043

ID ABV94043 standard; cDNA; 1429 BP.

XX AC ABV94043;

XX 08-JAN-2003 (first entry)

DE Breast carcinoma related nucleotide sequence SEQ ID NO:34.

DE Human; breast carcinoma; cancer; tumour; cytostatic; anti-tumour; gene;

XX ss.

XX Homo sapiens.

XX WO200246467-A2.

XX PD 13-JUN-2002.

XX 07-DEC-2001; 2001WO-IB002811.

XX 08-DEC-2000; 2000US-0254090P.

XX 07-DEC-2001; 2001US-00007926.

XX (IPSO-) IPSOGEN.

XX Bertucci F, Houlgatte R, Birnbaum D, Nguyen C, Viens P, Fert V;

XX WPI; 2002-619023/66.

XX Novel polynucleotide library useful in molecular characterization of a

XX carcinoma, comprising a pool of polynucleotide sequences or its

XX subences which are either underexpressed or overexpressed in tumor

XX cells.

XX Claim 1; Page 130-131; 401pp; English.

XX The present invention describes a polynucleotide library (I) useful in

XX the molecular characterisation of a carcinoma, comprising a pool of

XX polynucleotides or its subsequences which are either underexpressed or

XX overexpressed in tumour cells, and correspond to any of the

XX polynucleotide sequences chosen from the 468 sequences given in ABV94010

XX to ABV94477. Also described: (i) a polynucleotide array (II) useful for

XX the prognosis or diagnostic of tumour, comprising (I); and (2) detecting

XX (M1) differentially expressed polynucleotide sequences which are

XX correlated with a cancer, involves obtaining a polynucleotide sample from

XX a patient, and reacting the polynucleotide sample obtained with a probe

XX immobilised on a solid support, where the probe comprises any combination

XX of the polynucleotide sequences of (I) or its expression products encoded

XX by polynucleotide sequences of (I), and detecting the reaction product.

XX (i) have cytostatic activities and can be used as anti-tumour agents. (i)

XX is useful in molecular characterisation of a carcinoma. (I) and (II) are

XX useful for the prognosis or diagnostic of tumour, in differentiating a

XX normal cell from a cancer cell, detecting a hormone sensitive tumour

XX cell, differentiating a tumour with lymph nodes from a tumour without

XX lymph nodes, differentiating antitumour-sensitive tumours from

XX antitumour-insensitive tumours, and classifying good and poor prognosis

XX primary breast tumours. (i) is useful for large-scale molecular

XX characterisation of breast cancer that help in prediction, prognosis and

XX cancer treatment, and for detecting differentially expressed genes that

XX correlated with a cancer

XX SQ Sequence 1429 BP; 372 A; 384 C; 381 G; 292 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 1429;

Best Local Similarity 100.0%; Pred. No. 1.7e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

DB 1336 AAATGAAAAA 1384

RESULT 46

AAH72649

ID AAH72649 standard; cDNA; 1676 BP.

XX AC AAH72649;

XX 19-SEP-2001 (first entry)

XX Human cervical cancer marker nucleic acid 3923.

XX Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.

XX Homo sapiens.

XX WO200142467-A2.

XX PD 14-JUN-2001.

XX 08-DEC-2000; 2000WO-US033312.

XX 08-DEC-1999; 99US-0169681P.

XX 21-DEC-1999; 99US-0171350P.

XX 14-MAR-2000; 2000US-0189315P.

XX 12-MAY-2000; 2000US-0203791P.

XX 09-JUN-2000; 2000US-0210600P.

XX 21-JUL-2000; 2000US-0220114P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Deeds J, Berger A, Zhao X;

XX WPI; 2001-375006/39.

XX New isolated nucleic acid for diagnosing and treating cervical cancer and

XX for assessing and detecting compounds for treating the cancer.

XX Claim 1; Page 780-781; 1051pp; English.

XX The invention relates to novel genes (AAH68727-AAH73383) associated with
 CC cervical cancer with cytostatic activity. The nucleic acids and encoded
 CC polypeptides are useful; to assess if a patient is afflicted with
 CC cervical cancer or has a pre-malignant condition; to monitor the
 CC progression of cervical cancer or a premalignant condition in a patient;
 CC and to select and/or assess the efficacy of a compound or therapy for
 CC inhibiting cervical cancer in a patient. The nucleic acids may also be
 CC useful for gene therapy

XX Sequence 1676 BP; 556 A; 302 C; 380 G; 438 T; 0 U; 0 Other;
 SQ

Query Match 1.3%; Score 49; DB 4; Length 1676;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
 Db 1295 AAATGAAAAA 1343

RESULT 47
 ABV27467
 ID ABV27467 standard; cDNA; 1702 BP.
 XX AC ABV27467;
 XX DT 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 27458.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 5602; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 1702 BP; 560 A; 309 C; 393 G; 440 T; 0 U; 0 Other;
 XX Query Match 1.3%; Score 49; DB 5; Length 1702;
 XX Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
 Db 1321 AAATGAAAAA 1369

RESULT 48
 ABV27724
 ID ABV27724 standard; cDNA; 1702 BP.
 XX AC ABV27724;
 XX DT 16-SEP-2002 (first entry)
 XX DE Human prostate expression marker cDNA 27715.
 XX KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
 XX KW pharmacogenomic marker; gene; ss.
 XX OS Homo sapiens.
 XX PN WO200160860-A2.
 XX PD 23-AUG-2001.
 XX PF 20-FEB-2001; 2001WO-US005171.
 XX PR 17-FEB-2000; 2000US-0183319P.
 XX PR 16-MAR-2000; 2000US-0189862P.
 XX PR 25-MAY-2000; 2000US-0207454P.
 XX PR 09-JUN-2000; 2000US-0211314P.
 XX PR 18-JUL-2000; 2000US-0219007P.
 XX PR 13-DEC-2000; 2000US-0255281P.
 XX PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
 XX PI Schlegel R, Endege WO, Monahan JE;
 XX WPI; 2001-662795/76.
 XX Novel isolated nucleic acid molecule associated with cancerous state of
 PT prostate cells and correlating with presence of prostate cancer, useful
 PT for detecting presence of prostate cancer, stage of prostate cancer.
 XX Claim 1; Page 5689-5692; 11750pp; English.
 XX The invention relates to an isolated nucleic acid molecule (I) comprising
 CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
 CC specification or its complement. (I) is useful for: (a) assessing whether
 CC a patient is afflicted with prostate cancer; (b) monitoring the
 CC progression of prostate cancer in a patient; (c) assessing the efficacy
 CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
 CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
 CC (e) selecting a composition for inhibiting prostate cancer in a patient;
 CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
 CC determining whether prostate cancer has metastasized in a patient; (h)
 CC assessing the aggressiveness or indolence of prostate cancer in a patient
 CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX SQ Sequence 1702 BP; 560 A; 309 C; 393 G; 440 T; 0 U; 0 Other;
 XX Query Match 1.3%; Score 49; DB 5; Length 1702;
 XX Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762

Db 1321 AATGAAAAA 1369

RESULT 49

ABV25281

ID ABV25281 standard; cDNA; 1702 BP.

XX AC ABV25281;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 25272.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 4967; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 1702 BP; 560 A; 309 C; 393 G; 440 T; 0 U; 0 Other;

Query Match

Best Local Similarity 1.3%; Score 49; DB 5; Length 1702;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

Db 1321 AATGAAAAA 1369

RESULT 50

ABV21893

ID ABV21893 standard; cDNA; 1702 BP.

XX AC ABV21893;

DT 13-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 21884.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;

XX pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 3725; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (I) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 1702 BP; 560 A; 309 C; 393 G; 440 T; 0 U; 0 Other;

Query Match

Best Local Similarity 1.3%; Score 49; DB 5; Length 1702;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA 3762

Db 1321 AATGAAAAA 1369

RESULT 51

AAD07369

ID AAD07369 standard; DNA; 1933 BP.

XX AC AAD07369;

XX 10-AUG-2001 (first entry)

XX Human DNA encoding SAP-1 protein (marker 20).

XX Human; cytostatic; diagnosis; prostate disorder; prostate cancer; BPH; benign prostatic hyperplasia; SAP-1 protein; ds.

XX Homo sapiens.

XX PD 02-AUG-2001.
XX PF 17-JAN-2001; 2001WO-US001325.
XX PF 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-MAR-2000; 2000US-0198113P.
PR 19-MAY-2000; 2000US-0205535P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225268P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225575P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 23-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 06-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 02-OCT-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 13-OCT-2000; 2000US-0239337P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 08-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX FA Rosen CA, Barash SC, Ruben SM;
XX PI
XX XX

DR WPI; 2001-451925/48.
DR P-PSDB; AAE09698.
XX Isolated polypeptide for treating, preventing and/ or prognosing medical
PT disorders and also for testing and detection e.g. diagnosis and screening
PT for agonists.
XX Claim 1; SEQ ID NO 19; 469pp; English.
XX
XX AAD16750-AAD16775 represent cDNAs corresponding to novel human protein
CC genes, and AAE09690-AA09715 represent the proteins they encode. AAD16777
CC -AAD16780 represent novel human genomic DNA fragments. The novel proteins
CC and their DNAs are useful for diagnosing, treating, preventing and/or
CC prognosing inflammatory disorders (bursitis or tendonitis); neural
CC disorders (e.g. Parkinson's disease, Alzheimer's disease); immune system
CC disorders; AIDS; autoimmune diseases (e.g., rheumatoid arthritis);
CC muscular disorders; reproductive disorders; gastrointestinal disorders
CC (malabsorption syndrome, Crohn's disease); pulmonary disorders;
CC cardiovascular disorders (myocardial infarction, ischaemia, arrhythmias);
CC renal disorders (glomerulonephritis, nephrotic syndrome); cancerous
CC disease and conditions (breast cancer); hyperproliferative disorders
CC (leukaemia, hyperplasia); tumours; foetal and developmental abnormalities
CC ; haematopoietic disorders; respiratory disorders (rhinitis, asthma);
CC angiotensin disorders; diabetes; atherosclerosis; endocrine disorders;
CC pregnancy-related disorders and infections. The novel protein DNA is
CC useful in gene therapy and anti-sense therapy. The proteins can also be
CC used to aid wound healing and epithelial cell proliferation, to prevent
CC skin aging due to sunburn, to maintain organs before transplantation, for

Query Match 1.3%; Score 49; DB 4; Length 1998;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3714
DB 1906 AAATGAAAAA 1906

RESULT 54
ADC22050
ID ADC22050 standard; cDNA; 1998 BP.
AC ADC22050;
XX
XX 18-DEC-2003 (first entry)
DE Human cDNA from secreted protein gene 9 #1.
XX Secreted protein; cytostatic; antibacterial; virucide; neuroprotective;
KW synecological; gastrointestinal-gen; cardiant; cardiovascular-gen;
KW nephrotropic; antiinflammatory; muscular-gen; respiratory-gen;
KW immunosuppressive; cerebroprotective; vasotropic; nootropic;
KW antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; Human; ss; gene.
XX
XX Homo sapiens.
XX
XX US2003082681-A1.
XX
XX 01-MAY-2003.
XX
XX 07-MAR-2002; 2002US-00091391.
XX
XX 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.

19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214866P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.
PR 14-AUG-2000; 2000US-0225269P.
PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226686P.
PR 22-AUG-2000; 2000US-0227183P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 06-SEP-2000; 2000US-0230438P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232397P.
PR 14-SEP-2000; 2000US-0232398P.
PR 14-SEP-2000; 2000US-0232399P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234224P.
PR 25-SEP-2000; 2000US-0234977P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235634P.
PR 27-SEP-2000; 2000US-0235636P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.
PR 17-JAN-2001; 2001US-00764903.
PA (HUMA-) HUMAN GENOME SCI INC.
XX
XX
PI Rosen CA, Ruben SM, Barash SC;
XX
XX WPI; 2003-786903/74.
DR P-PSDB; ADC22076.
XX
XX New nucleic acid molecules and polypeptides for diagnosing, preventing or
PT treating disorders associated with aberrant expression of the
PT polypeptide, e.g. neural or cardiovascular disorders, and in chromosome
PT identification.
XX
XX Claim 1; SEQ ID NO 19; 242pp; English.
PS
XX The invention relates to an isolated nucleic acid molecule (cDNA)
CC

CC encoding a human secreted protein, representing one of 15 novel genes.
CC Also included are recombinant vectors, host cells (expressing the
CC protein), the secreted proteins (including their fragments, epitopes and
CC homologues), an isolated antibody that binds specifically to the protein,
CC diagnosing a pathological condition or susceptibility to a pathological
CC condition (comprising determining the presence or absence of a mutation
CC in the nucleic acid and diagnosing a condition based on the presence or
CC absence of the mutation), diagnosing a pathological condition or
CC susceptibility to a pathological condition (comprising determining the
CC presence or amount of expression of the protein in a biological sample
CC and diagnosing a condition based on the presence or amount of expression
CC of the protein), preventing, treating or ameliorating a medical condition
CC by administering the nucleic acid or protein to a mammalian subject,
CC identifying a binding partner to the protein, the gene corresponding to
CC the cDNA sequence, and identifying an activity in a biological assay
CC (comprising expressing the nucleic acid in a cell, isolating the
CC supernatant, detecting an activity in a biological assay and identifying
CC the protein in the supernatant having the activity). The nucleic acids
CC and proteins display the following activities: Cytostatic, antibacterial,
CC Virucide, Neuroprotective, Gynaecological, Gastrointestinal-Gen,
CC Cardiant, Cardiovascular-Gen, Nephrotropic, Antiinflammatory, Muscular-

Query Match 1.3%; Score 49; DB 9; Length 1998;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1906 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1954

RESULT 55
ACD13422
ID ACD13422 standard; cDNA; 2038 BP.
XX AC
XX ACD13422;
XX
DT 13-AUG-2003 (first entry)
XX
DE Human DNA encoding a p53 modifier, SEQ ID 94.
XX
XX Human; ss; gene; p53 modifier; cytostatic; cancer; cytostatic;
KW antiangiogenic; antiapoptotic; p53 pathway; breast cancer; colon cancer;
KW kidney cancer; lung cancer; ovarian cancer; angiogenesis; cell cycle;
KW apoptotic disorder; cell proliferation disorder.
XX
OS Homo sapiens.
XX
XX WO200299122-A1.
PN
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017382.
XX
XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX
XX (EXEL-) EXELIXIS INC.
XX
XX Friedman L, Flowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-156859/15.
DR P-PSDB; ABO07249.
XX
XX Identifying modulators of the p53 pathway for use in treating apoptotic
PT or cell proliferation disorders, comprises screening for agents that
PT modulate activity of a human ortholog of genes that modify the p53
PT pathway in Drosophila.
XX
XX Example 2; Page 376-377; 678pp; English.
PS
XX The invention relates to identifying (M1) a candidate p53 pathway

modulating agent, by contacting an assay system comprising a purified HM polypeptide (human orthologue of genes that modify the p53 pathway in Drosophila) or nucleic acid with a test agent under conditions, where but for the presence of the test agent, the system provides a reference activity, and detecting a test-agent-biased activity of the assay system. Also included are modulating (M2) a p53 pathway of a cell (comprising contacting a cell defective in p53 function with a candidate modulator that specifically binds to a HM polypeptide comprising an HM amino acid sequence, where p53 function is restored), modulating (M3) a p53 pathway in a mammalian cell (comprising contacting the cell with an agent that specifically binds an HM polypeptide or nucleic acid) and diagnosing (M4) a disease in a patient (comprising: (a) obtaining a biological sample from the patient; (b) contacting the sample with a probe for HM expression; (c) comparing the results with a control; and (d) determining whether the comparison indicates a likelihood disease). (M1) is useful for identifying modulators of the p53 pathway. A probe for HM expression is useful for diagnosing breast, colon, kidney, lung and ovarian cancer, in a patient, where the cancer has greater than 25 % expression level. Modulators identified by (M1) are useful in a variety of diagnostic and therapeutic applications, where disease or disorder prognosis is related to defects in the p53 pathway, such as, angiogenesis, apoptotic or cell proliferation disorders (e.g. cancer). Another two new methods (M2 and M3) are useful for modulating the p53 pathway of a cell, thus restoring the p53 function of the cell, so that the cell undergoes normal proliferation or progression through the cell cycle. (M2) and (M3) are also useful for treating defects in the p53 pathway such as angiogenic, apoptotic or cell proliferation disorders. The present sequence is an HM nucleic acid encoding a p53 pathway modifying protein

Sequence 2038 BP; 527 A; 580 C; 567 G; 364 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 7; Length 2038;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 1974 AAATGAAAAA 2022

RESULT 56

ABZ71965
ID ABZ71965 standard; cDNA; 2106 BP.

AC ABZ71965;

DT 01-APR-2003 (first entry)

XX Human cDNA for transketolase GenBank X67688.

XX Human; cancer; stomach cancer; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200283899-A1.

XX 24-OCT-2002.

XX 28-MAR-2002; 2002WO-JP003038.

XX 10-APR-2001; 2001JP-00112039.

XX 21-SEP-2001; 2001JP-00280193.

XX (TAKA-) TAKARA BIO INC.

XX Yoshikawa Y, Okamoto S, Oura T, Mineno J, Asada K, Kato I;
PI Inoue H, Mori M;

XX WPI; 2003-093022/08.

XX Measuring changes in expression of 264 cancer associated genes for
PT detection of stomach cancer and screening of potential anticancer agents.

PS Claim 2; Page; 266pp; Japanese.

XX The invention relates to a method for the detection of cancer in which a change in the expression of 1 or more of 264 specified cancer associated genes, ABZ71694-ABZ71957, or of sequences at least 80% homologous to them in the specimen tissue as compared to normal tissue is observed. The genes are used in detection, diagnosis and treatment of cancer, especially of stomach cancer. The present sequence is that of a cancer associated polynucleotide of the invention. Note: The present sequence was not given in the printed specification but was isolated using the GenBank accession number given in the DE line

XX Sequence 2106 BP; 529 A; 608 C; 589 G; 380 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 7; Length 2106;
Best Local Similarity 100.0%; Pred. No. 1.6e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 3762
DB 2044 AAATGAAAAA 2092

RESULT 57

ABQ54860/c

ID ABQ54860 standard; cDNA; 2143 BP.

XX ABQ54860;

XX 22-AUG-2002 (first entry)

XX Human ovarian antigen HAZAA31 cDNA, SEQ ID NO:740.

XX Human; ovarian antigen; ovary; ovarian; breast; cancer; tumour;
XX ovarian cancer; breast cancer; tumour; reproductive system disorder;
XX infertility; pregnancy disorder; anovulation; polycystic ovary syndrome;
XX PCOS; ovarian cyst; dysmenorrhea; endocrine disorder; infection;
XX inflammatory condition; immune disorder; blood disorder;
XX cardiovascular disorder; respiratory disorder; neurological disorder;
XX gastrointestinal disorder; urinary system disorder; drug screening;
XX gene therapy; chromosome mapping; forensic analysis;
XX antibody preparation; cytostatic; immunomodulatory; neuroprotective;
XX antiinflammatory; gynaecological; reproductive; chromosome 3p14.3; gene; ss.

XX Homo sapiens.

XX WO200200677-A1.

XX 03-JAN-2002.

XX 07-JUN-2001; 2001WO-US018569.

XX 07-JUN-2000; 2000US-0209467P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Birse CE, Rosen CA;

XX WPI; 2002-147878/19.

XX P-PSDB; ABP41783.

XX Isolated nucleic acid molecules encoding novel ovarian polypeptides,
XX useful in the prevention, treatment and diagnosis of cancer (e.g. ovarian
XX cancer), immune disorders, cardiovascular disorders and neurological
XX diseases.

XX Claim 1; SEQ ID NO 740; 2922pp; English.

XX The invention relates to 2175 novel human ovarian antigens (ABP41054-
XX ABP43228) and to cDNAs encoding them (ABQ54311-ABQ56305), and also
XX encompasses polypeptides 90% identical and polynucleotides 95% identical
XX to the sequences of the invention. The invention additionally relates to

CC recombinant vectors and host cells comprising human ovarian antigen
 CC polynucleotides, antibodies against human ovarian antigens, and the use
 CC of ovarian antigen polynucleotides and polypeptides in diagnosing,
 CC treating, prognosing or preventing various ovary and/or breast-related
 CC disorders. Such conditions include ovarian cancer and breast cancer, and
 CC metastatic tumours of ovarian or breast origin, reproductive system
 CC disorders (e.g., infertility, disorders of pregnancy, anovulation,
 CC polycystic ovary syndrome, ovarian cysts, and dysmenorrhoea), endocrine
 CC disorders, infections (e.g., chlamydia, HIV, toxoplasmosis, and toxic
 CC shock syndrome), inflammatory conditions (e.g., mastitis, oophoritis and
 CC vaginitis), immune disorders (e.g., congenital and acquired
 CC immunodeficiencies, autoimmune oophoritis, systemic lupus erythematosus),
 CC blood-related disorders (e.g., anaemia), cardiovascular disorders,
 CC respiratory disorders, neurological disorders, gastrointestinal disorders
 CC and urinary system disorders. Ovarian antigen polypeptides and
 CC polynucleotides may also be used in screening for compounds which
 CC modulate ovarian antigen expression or activity. The polynucleotides may
 CC further be used for gene therapy, chromosome mapping, in the
 CC identification of individuals and in forensic analysis, and the
 CC polypeptides may be used as food additives or to prepare antibodies
 CC useful in disease diagnosis, drug targeting and phenotyping. The present
 CC sequence represents cDNA encoding a human ovarian antigen of the
 CC invention. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX SQ Sequence 2143 BP; 377 A; 598 C; 620 G; 545 T; 0 U; 3 Other;

Query Match 1.3%; Score 49; DB 6; Length 2143;
 Best Local Similarity 100.0%; Pred. No. 1.6e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA AA 3762

DB 80 AATGAAAAA AA 32

RESULT 58

AAC76644

ID AAC76644 standard; cDNA; 2543 BP.

XX AC AAC76644;

XX DT 08-FEB-2001 (first entry)

XX DE Human ORFX ORF2199 polynucleotide sequence SEQ ID NO:4397.

XX Human; open reading frame; ORFX; detection; cytostatic; hepatotropic;
 KW vulnery; antipsoriatic; antiparkinsonian; noctropic; neuroprotective;
 KW anticonvulsant; osteopathic; antiarthritic; immunosuppressant; cardiant;
 KW immunostimulant; thrombolytic; coagulant; vasotropic; antidiabetic;
 KW hypotensive; dermatological; immunosuppressive; antiinflammatory;
 KW antiviral; antibacterial; antifungal; antirheumatic; antithyroid;
 KW antianemic; gene therapy; cancer; proliferative disorder; hypertension;
 KW neurodegenerative disorder; osteoarthritis; graft vs host disease;
 KW cardiovascular disease; diabetes mellitus; hypothyroidism; SCID; AIDS;
 KW cholesterol ester storage; systemic lupus erythematosus; infection;
 KW severe combined immunodeficiency; malaria; autoimmune disorder; asthma;
 KW allergy; aplastic anaemia; nocturnal haemoglobinuria; burn; wound;
 KW bone damage; cartilage damage; antiinflammatory disease; coagulation;
 KW thrombosis; contraceptive; ss.

XX Homo sapiens.

XX WO200058473-A2.

XX PD 05-OCT-2000.

XX 31-MAR-2000; 2000WO-US008621.

XX 31-MAR-1999; 99US-0127607P.

XX 02-APR-1999; 99US-0127636P.

XX 05-APR-1999; 99US-0127728P.

PR 30-MAR-2000; 2000US-00540763.

XX (CURA-) CURAGEN CORP.

XX PI Shimkets RA, Leach M;

XX WPI; 2000-602362/57.

XX P-PSDB; AAB42435.

XX Novel nucleic acids and peptides derived from open reading frame X,
 XX useful for treating e.g. cancers, proliferative disorders,
 XX neurodegenerative disorders and cardiovascular disease.

XX Claim 5; Page 3587-3589; 5507pp; English.

XX AAC74446 to AAC77606 encode the proteins given in AAB40237 to AAB43397,
 XX which represent the human ORFX open reading frames 1 to 3161. The ORFX
 XX sequences have activities such as: cytostatic; hepatotropic; vulnery;
 XX antipsoriatic; antiparkinsonian; noctropic; neuroprotective; osteopathic;
 XX anticonvulsant; antiarthritic; immunosuppressant; immunostimulant;
 XX cardiant; thrombolytic; coagulant; vasotropic; antidiabetic; hypotensive;
 XX dermatological; immunosuppressive; antiinflammatory; antibacterial;
 XX antiviral; antifungal; antirheumatic; antithyroid; and antianemic. The
 XX sequences can be used for determining the presence of or predisposition
 XX to, or preventing or treating pathological conditions associated with an
 XX ORFX-associated disorder. The nucleic acids can be used to express ORFX
 XX proteins in gene therapy vectors. The proteins and nucleic acids may be
 XX used to treat cancers, proliferative disorders, neurodegenerative
 XX disorders, osteoarthritis, graft vs host disease, cardiovascular disease,
 XX diabetes mellitus, hypertension, hypothyroidism, cholesterol ester
 XX storage, systemic lupus erythematosus, severe combined immunodeficiency
 XX (SCID), AIDS, viral, bacterial or fungal infection, malaria, autoimmune
 XX disorders, asthma, allergies, aplastic anaemia, burns, wounds, bone and
 XX cartilage damage, nocturnal haemoglobinuria, antiinflammatory disease; to
 XX enhance coagulation; to inhibit thrombosis; and as a contraceptive

XX SQ Sequence 2543 BP; 630 A; 589 C; 684 G; 639 T; 0 U; 1 Other;

Query Match 1.3%; Score 49; DB 3; Length 2543;

Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

DB 2494 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 2542

RESULT 59

AAD43556

ID AAD43556 standard; cDNA; 2713 BP.

XX AC AAD43556;

XX DT 14-NOV-2002 (first entry)

XX DE Human CD2000 cDNA.

XX Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein;
 KW immune proliferative disorder; immune disorder; rheumatoid arthritis;
 KW carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease;
 KW Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic;
 KW osteoarthritis; allergic inflammatory disorder; viral infection; asthma;
 KW psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis;
 KW diabetes mellitus; septic shock; chronic obstructive pulmonary disease;
 KW emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis;
 KW acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic;
 KW immunosuppressive; neuroprotective; antiinflammatory; Crohn's disease;
 KW osteopathic; antibacterial; immunomodulator; inflammatory bowel disease;
 KW jaundice; dermatological; ulcerative colitis; AIDS; chromosome 1q23;
 KW Gene; ss.

XX Homo sapiens.

```

FH Key Location/Qualifiers
FT CDS 64..1059
FT     /*tag= a
FT     /product= "Human CD2000 protein"
FT     64..1056
FT     /*tag= d
FT     /notes= "This region designated as SEQ.ID.NO:2 is
FT     specifically referred in claim 2"
FT     64..129
FT     /*tag= b
FT     mat_peptide 130..1056
FT     /*tag= c
FT     /product= "Human mature CD2000 protein"
XX
FN EP1223218-A1.
XX
XX 17-JUL-2002.
XX
XX 02-NOV-2001; 2001EP-00309339.
XX
XX 03-NOV-2000; 2000US-00706167.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Fraser CC;
XX
XX WPI; 2002-620680/67.
XX
XX P-FSDB; AAE28220.
XX
XX Novel isolated polypeptide containing immunoglobulin and immunoglobulin-
XX like domains and SLAM associated protein, termed CD2000 or CD2001, useful
XX for treating immune, inflammatory, or hepatic circulatory disorders.
XX
XX Claim 1; Page 66; 138pp; English.
XX
XX The invention relates to nucleic acid molecule, designated CD2000 which
XX encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains
XX and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is
XX useful for treating disorder such as immune proliferative disorders,
XX immune disorders (e.g. carcinoma), viral infection, autoimmune disorders
XX (e.g., arthritis, multiple sclerosis, Grave's disease, and Hashimoto's
XX disease), T cell disorder (e.g. acquired immune deficiency syndrome
XX (AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative
XX colitis), inflammatory disorders (e.g. rheumatoid arthritis and
XX osteoarthritis), allergic inflammatory disorders (e.g. systemic lupus erythematosus, and
XX insulin-dependent diabetes mellitus), cytotoxic disorders, septic shock,
XX psoriasis), apoptotic disorders (e.g. systemic lupus erythematosus, and
XX chronic obstructive pulmonary disease (e.g. emphysema), bronchitis,
XX cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis,
XX acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in
XX gene therapy. CD2000 DNA is useful in screening assays, detection assays
XX (e.g. chromosomal mapping, tissue typing, forensic biology), predictive
XX medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
XX trials and pharmacogenomics), and in methods of treatment (e.g.
XX therapeutic and prophylactic). The present sequence is human CD2000 cDNA.
XX
XX Human CD2000 gene is located at chromosome 1q23
XX
XX Sequence 2713 BP; 799 A; 647 C; 529 G; 738 T; 0 U; 0 Other;
XX
XX Query Match 1.3%; Score 49; DB 6; Length 2713;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-07;
XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 3714 AAATGAAAAA 3762
XX 2650 AAATGAAAAA 2698
XX
XX RESULT 60
XX AAD43566
XX AAD43566 standard; DNA; 2713 BP.
XX
XX AAD43566;

```

```

XX 14-NOV-2002 (first entry)
XX Human CD2000 DNA #3.
XX
XX Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein;
XX immune proliferative disorder; immune disorder; rheumatoid arthritis;
XX carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease;
XX Hashimoto's disease; acquired immune deficiency syndrome; hepatocellular;
XX osteoarthritis; allergic inflammatory disorder; viral infection; asthma;
XX psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis;
XX diabetes mellitus; septic shock; chronic obstructive pulmonary disease;
XX emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis;
XX acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic;
XX immunosuppressive; neuroprotective; antiinflammatory; Crohn's disease;
XX osteopathic; antibacterial; immunomodulator; inflammatory bowel disease;
XX jaundice; dermatological; ulcerative colitis; AIDS; gene; ds.
XX
XX Homo sapiens.
XX
XX EP1223218-A1.
XX
XX 17-JUL-2002.
XX
XX 02-NOV-2001; 2001EP-00309339.
XX
XX 03-NOV-2000; 2000US-00706167.
XX
XX (MILL-) MILLENNIUM PHARM INC.
XX
XX Fraser CC;
XX
XX WPI; 2002-620680/67.
XX
XX Novel isolated polypeptide containing immunoglobulin and immunoglobulin-
XX like domains and SLAM associated protein, termed CD2000 or CD2001, useful
XX for treating immune, inflammatory, or hepatic circulatory disorders.
XX
XX Disclosure; Page 79; 138pp; English.
XX
XX The invention relates to nucleic acid molecule, designated CD2000 which
XX encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains
XX and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is
XX useful for treating disorder such as immune proliferative disorders,
XX immune disorders (e.g. carcinoma), viral infection, autoimmune disorders
XX (e.g., arthritis, multiple sclerosis, Grave's disease, and Hashimoto's
XX disease), T cell disorder (e.g. acquired immune deficiency syndrome
XX (AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative
XX colitis), inflammatory disorders (e.g. rheumatoid arthritis and
XX osteoarthritis), allergic inflammatory disorders (e.g. systemic lupus erythematosus, and
XX insulin-dependent diabetes mellitus), cytotoxic disorders, septic shock,
XX chronic obstructive pulmonary disease (e.g. emphysema), bronchitis,
XX cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis,
XX acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in
XX gene therapy. CD2000 DNA is useful in screening assays, detection assays
XX (e.g. chromosomal mapping, tissue typing, forensic biology), predictive
XX medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
XX trials and pharmacogenomics), and in methods of treatment (e.g.
XX therapeutic and prophylactic). The present sequence is human CD2000 DNA.
XX
XX Human CD2000 gene is located at chromosome 1q23
XX
XX Sequence 2713 BP; 799 A; 646 C; 529 G; 739 T; 0 U; 0 Other;
XX
XX Query Match 1.3%; Score 49; DB 6; Length 2713;
XX Best Local Similarity 100.0%; Pred. No. 1.5e-07;
XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX
XX 3714 AAATGAAAAA 3762
XX 2650 AAATGAAAAA 2698
XX
XX RESULT 61

```


AAD43565
 ID AAD43565 standard; DNA; 2713 BP.
 AC AAD43565;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Human CD2000 DNA #2.
 XX
 KW Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein;
 KW immune proliferative disorder; immune disorder; rheumatoid arthritis;
 KW carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease;
 KW Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic;
 KW osteoarthritis; allergic inflammatory disorder; viral infection; asthma;
 KW psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis;
 KW diabetes mellitus; septic shock; chronic obstructive pulmonary disease;
 KW emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis;
 KW acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic;
 KW immunosuppressive; neuroprotective; antiinflammatory; Crohn's disease;
 KW osteopathic; antibacterial; immunomodulator; inflammatory bowel disease;
 KW jaundice; dermatological; ulcerative colitis; AIDS; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN EP1223218-A1.
 XX
 PD 17-JUL-2002.
 XX
 PF 02-NOV-2001; 2001EP-00309339.
 XX
 PR 03-NOV-2000; 2000US-00706167.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Fraser CC;
 XX
 DR WPI; 2002-620680/67.
 XX
 PT Novel isolated polypeptide containing immunoglobulin and immunoglobulin-
 PT like domains and SLAM associated protein, termed CD2000 or CD2001, useful
 PT for treating immune, inflammatory, or hepatic circulatory disorders.
 XX
 PS Disclosure; Page 78; 138pp; English.
 XX
 CC The invention relates to nucleic acid molecule, designated CD2000 which
 CC encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains
 CC and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is
 CC useful for treating disorder such as immune proliferative disorders,
 CC immune disorders (e.g. carcinoma), viral infection, autoimmune disorders
 CC (e.g. arthritis, multiple sclerosis, Grave's disease, and Hashimoto's
 CC disease), T cell disorder (e.g. acquired immune deficiency syndrome
 CC (AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative
 CC colitis), inflammatory disorders (e.g. rheumatoid arthritis and
 CC osteoarthritis), allergic inflammatory disorders (e.g. emphysema), bronchitis,
 CC cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis,
 CC acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in
 CC gene therapy. CD2000 DNA is useful in screening assays, detection assays
 CC (e.g. chromosomal mapping, tissue typing, forensic biology), predictive
 CC medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
 CC trials and pharmacogenomics), and in methods of treatment (e.g.
 CC therapeutic and prophylactic). The present sequence is human CD2000 DNA
 XX
 SQ Sequence 2713 BP; 798 A; 648 C; 529 G; 738 T; 0 U; 0 Other;
 Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;
 Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AATCGAAAAA
 |||||

Db 2650 AATCGAAAAA
 RESULT 62
 AAD43564
 ID AAD43564 standard; DNA; 2713 BP.
 XX
 AC AAD43564;
 XX
 DT 14-NOV-2002 (first entry)
 XX
 DE Human CD2000 DNA #1.
 XX
 KW Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein;
 KW immune proliferative disorder; immune disorder; rheumatoid arthritis;
 KW carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease;
 KW Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic;
 KW osteoarthritis; allergic inflammatory disorder; viral infection; asthma;
 KW psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis;
 KW diabetes mellitus; septic shock; chronic obstructive pulmonary disease;
 KW emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis;
 KW acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic;
 KW immunosuppressive; neuroprotective; antiinflammatory; Crohn's disease;
 KW osteopathic; antibacterial; immunomodulator; inflammatory bowel disease;
 KW jaundice; dermatological; ulcerative colitis; AIDS; gene; ds.
 XX
 OS Homo sapiens.
 XX
 PN EP1223218-A1.
 XX
 PD 17-JUL-2002.
 XX
 PF 02-NOV-2001; 2001EP-00309339.
 XX
 PR 03-NOV-2000; 2000US-00706167.
 XX
 PA (MILL-) MILLENNIUM PHARM INC.
 XX
 PI Fraser CC;
 XX
 DR WPI; 2002-620680/67.
 XX
 PT Novel isolated polypeptide containing immunoglobulin and immunoglobulin-
 PT like domains and SLAM associated protein, termed CD2000 or CD2001, useful
 PT for treating immune, inflammatory, or hepatic circulatory disorders.
 XX
 PS Disclosure; Page 77-78; 138pp; English.
 XX
 CC The invention relates to nucleic acid molecule, designated CD2000 which
 CC encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains
 CC and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is
 CC useful for treating disorder such as immune proliferative disorders,
 CC immune disorders (e.g. carcinoma), viral infection, autoimmune disorders
 CC (e.g. arthritis, multiple sclerosis, Grave's disease, and Hashimoto's
 CC disease), T cell disorder (e.g. acquired immune deficiency syndrome
 CC (AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative
 CC colitis), inflammatory disorders (e.g. rheumatoid arthritis and
 CC osteoarthritis), allergic inflammatory disorders (e.g. emphysema), bronchitis,
 CC cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis,
 CC acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in
 CC gene therapy. CD2000 DNA is useful in screening assays, detection assays
 CC (e.g. chromosomal mapping, tissue typing, forensic biology), predictive
 CC medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
 CC trials and pharmacogenomics), and in methods of treatment (e.g.
 CC therapeutic and prophylactic). The present sequence is human CD2000 DNA
 XX
 SQ Sequence 2713 BP; 799 A; 646 C; 529 G; 739 T; 0 U; 0 Other;
 Query Match 1.3%; Score 49; DB 6; Length 2713;
 Best Local Similarity 100.0%; Pred. No. 1.5e-07;

Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2650 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 63

AAAD3567
ID AAD43567 standard; DNA; 2713 BP.

XX
AC AAD43567;

XX
DT 14-NOV-2002 (first entry)

XX
DE Human CD2000 DNA #4.

XX
Human; immunoglobulin; Ig; SLAM associated protein; SAP; CD2000 protein;
immune proliferative disorder; immune disorder; rheumatoid arthritis;
carcinoma; autoimmune disorder; multiple sclerosis; Grave's disease;
Hashimoto's disease; acquired immune deficiency syndrome; hepatotropic;
osteoarthritis; allergic inflammatory disorder; viral infection; asthma;
psoriasis; apoptotic disorder; systemic lupus erythematosus; bronchitis;
diabetes mellitus; septic shock; chronic obstructive pulmonary disease;
emphysema; cachexia; hepatic circulatory disorder; hepatitis; cirrhosis;
acute myeloid leukaemia; haemophilia; anaemia; gene therapy; cytostatic;
immunosuppressive; neutropenic; antineoplastic; Crohn's disease;
osteopathic; antibacterial; immunomodulator; inflammatory bowel disease;
jaundice; dermatological; ulcerative colitis; AIDS; gene; ds.

XX
OS Homo sapiens.

XX
FN EP1223218-A1.

XX
PD 17-JUL-2002.

XX
PF 02-NOV-2001; 2001EP-00309339.

XX
PR 03-NOV-2000; 2000US-00706167.

XX
PA (MILL-) MILLENNIUM PHARM INC.

XX
PI Fraser CC;

XX
DR WPI; 2002-620680/67.

XX
Novel isolated polypeptide containing immunoglobulin and immunoglobulin-
like domains and SLAM associated protein, termed CD2000 or CD2001, useful
for treating immune, inflammatory, or hepatic circulatory disorders.

XX
PS Disclosure; Page 79-80; 139pp; English.

XX
The invention relates to nucleic acid molecule, designated CD2000 which
encodes a polypeptide containing immunoglobulin (Ig) and Ig-like domains
and SLAM associated protein (SAP) motifs. CD2000 DNA and protein is
useful for treating disorder such as immune proliferative disorders,
immune disorders (e.g. carcinoma), viral infection, autoimmune disorders
(e.g., arthritis, multiple sclerosis, Grave's disease, and Hashimoto's
disease), T cell disorder (e.g. acquired immune deficiency syndrome
(AIDS)), inflammatory bowel disease (e.g. Crohn's disease and ulcerative
colitis), inflammatory disorders (e.g. rheumatoid arthritis and
psoriasis), apoptotic disorders (e.g. systemic lupus erythematosus, and
osteoarthritis), allergic inflammatory disorders (e.g. asthma and
insulin-dependent diabetes mellitus), cytotoxic disorders, septic shock,
chronic obstructive pulmonary disease (e.g. emphysema), bronchitis,
cachexia, jaundice, hepatic circulatory disorders, hepatitis, cirrhosis,
acute myeloid leukaemia, haemophilia and anaemia. CD2000 DNA is used in
gene therapy. CD2000 DNA is useful in screening assays, detection assays
(e.g. chromosomal mapping, tissue typing, forensic biology), predictive
medicine (e.g. diagnostic assays, prognostic assays, monitoring clinical
trials and pharmacogenomics), and in methods of treatment (e.g.
therapeutic and prophylactic). The present sequence is human CD2000 DNA

SQ Sequence 2713 BP; 799 A; 648 C; 529 G; 737 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2650 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2713;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 64

ABN86559
ID ABN86559 standard; DNA; 2851 BP.

XX
AC ABN86559;

XX
DT 05-NOV-2002 (first entry)

XX
DE Canine low affinity IgE receptor (CD23) nCaCD23_2851 DNA sequence.

XX
KW Canine; immunoglobulin E; IgE; CD23; antiallergic; antiasthmatic; gene;
antiinflammatory; dermatological; gene therapy; vaccine; ds.

XX
OS Canis familiaris.

XX
FH Key Location/Qualifiers

XX
FT CDS 199..1077

XX
FT /tag= a

XX
FT /product= "CD23 receptor"

XX
PN US6410714-B1.

XX
XX 25-JUN-2002.

XX
XX 24-MAR-2000; 2000US-00535521.

XX
XX 24-MAR-1999; 99US-0125913P.

XX
XX (HESK-) HESKA CORP.

XX
XX Weber ER, McCall CA;

XX
XX WPI; 2002-588996/63.

XX
XX P-PSDB; ABB81056.

XX
New isolated canine low affinity immunoglobulin E receptor nucleic acid
molecule, useful for protecting canids from diseases mediated by the
receptor, such as allergy, atopic dermatitis, asthma, and hay fever.

XX
PS Claim 1a; Col 31-36; 33pp; English.

XX
The invention relates to isolated canine low affinity immunoglobulin E
(IgE) receptor (CD23) polypeptides and encoding nucleic acid molecules.
The CD23 polypeptides can be expressed by standard recombinant
methodology. The CD23 polynucleotides are useful for protecting canids
from diseases mediated by CD23, for developing compounds that regulate
IgE and/or CD23 levels in a canid for treating allergy related diseases
such as atopic dermatitis, asthma, hay fever and food sensitivities. The
present sequence represents a CD23 nucleic acid molecule nCaCD23_2851

XX
SQ Sequence 2851 BP; 604 A; 938 C; 781 G; 528 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Db 2802 AAATGAAAAA 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

RESULT 65

```
ABN86560/c
ID  ABN86560 standard; DNA; 2851 BP.
XX  AC
XX  ABN86560;
XX  DT
XX  05-NOV-2002 (first entry)
XX  DE
XX  Canine CD23 nCaCD23_2851 complementary DNA sequence.
XX  KW
XX  Canine; immunoglobulin E; IgE; CD23; antiallergic; antiasthmatic;
XX  antiinflammatory; dermatological; gene therapy; vaccine; ds.
XX  KW
XX  Canis familiaris.
XX  OS
XX  US6410714-B1.
XX  FN
XX  25-JUN-2002.
XX  PD
XX  24-MAR-2000; 2000US-00535521.
XX  PF
XX  24-MAR-1999; 99US-0125913P.
XX  PR
XX  (HESK-) HESKA CORP.
XX  FA
XX  Weber ER, McCall CA;
XX  PI
XX  WPI; 2002-588896/63.
XX  DR
XX  New isolated canine low affinity immunoglobulin E receptor nucleic acid
XX  PT molecule, useful for protecting canids from diseases mediated by the
XX  PT receptor, such as allergy, atopic dermatitis, asthma, and hay fever.
XX  PT
XX  Claim 1b; Col 37-40; 33pp; English.
XX  PS
XX  The invention relates to isolated canine low affinity immunoglobulin E
XX  CC (IgE) receptor (CD23) polypeptides and encoding nucleic acid molecules.
XX  CC The CD23 polypeptides can be expressed by standard recombinant
XX  CC methodology. The CD23 polynucleotides are useful for protecting canids
XX  CC from diseases mediated by CD23, for developing compounds that regulate
XX  CC IgE and/or CD23 levels in a canid for treating allergy related diseases
XX  CC such as atopic dermatitis, asthma, hay fever and food sensitivities. The
XX  CC present sequence represents the complementary sequence of the CD23
XX  CC nucleic acid molecule nCaCD23_2851
XX  CC
XX  SQ Sequence 2851 BP; 528 A; 781 C; 938 G; 604 T; 0 U; 0 Other;
Query Match 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3714 AAATGAAAAA 3762
Db 50 AAATGAAAAA 2
RESULT 66
ABT18794/c
ID  ABT18794 standard; DNA; 2944 BP.
XX  AC
XX  ABT18794;
XX  DT
XX  16-APR-2003 (first entry)
XX  DE
XX  Aspergillus fumigatus essential gene #1152.
XX  KW
XX  Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX  KW cancer; contamination; biofilm; antibody; immune response; ds.
XX  KW
XX  Aspergillus fumigatus.
XX  OS
XX  WO200286090-A2.
XX  FN
XX  31-OCT-2002.
XX  PD
```

```
ABN86560/c
ID  ABN86560 standard; DNA; 2851 BP.
XX  AC
XX  ABN86560;
XX  DT
XX  05-NOV-2002 (first entry)
XX  DE
XX  Canine CD23 nCaCD23_2851 complementary DNA sequence.
XX  KW
XX  Canine; immunoglobulin E; IgE; CD23; antiallergic; antiasthmatic;
XX  antiinflammatory; dermatological; gene therapy; vaccine; ds.
XX  KW
XX  Canis familiaris.
XX  OS
XX  US6410714-B1.
XX  FN
XX  25-JUN-2002.
XX  PD
XX  24-MAR-2000; 2000US-00535521.
XX  PF
XX  24-MAR-1999; 99US-0125913P.
XX  PR
XX  (HESK-) HESKA CORP.
XX  FA
XX  Weber ER, McCall CA;
XX  PI
XX  WPI; 2002-588896/63.
XX  DR
XX  New isolated canine low affinity immunoglobulin E receptor nucleic acid
XX  PT molecule, useful for protecting canids from diseases mediated by the
XX  PT receptor, such as allergy, atopic dermatitis, asthma, and hay fever.
XX  PT
XX  Claim 1b; Col 37-40; 33pp; English.
XX  PS
XX  The invention relates to isolated canine low affinity immunoglobulin E
XX  CC (IgE) receptor (CD23) polypeptides and encoding nucleic acid molecules.
XX  CC The CD23 polypeptides can be expressed by standard recombinant
XX  CC methodology. The CD23 polynucleotides are useful for protecting canids
XX  CC from diseases mediated by CD23, for developing compounds that regulate
XX  CC IgE and/or CD23 levels in a canid for treating allergy related diseases
XX  CC such as atopic dermatitis, asthma, hay fever and food sensitivities. The
XX  CC present sequence represents the complementary sequence of the CD23
XX  CC nucleic acid molecule nCaCD23_2851
XX  CC
XX  SQ Sequence 2851 BP; 528 A; 781 C; 938 G; 604 T; 0 U; 0 Other;
Query Match 1.3%; Score 49; DB 6; Length 2851;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3714 AAATGAAAAA 3762
Db 50 AAATGAAAAA 2
RESULT 66
ABT18794/c
ID  ABT18794 standard; DNA; 2944 BP.
XX  AC
XX  ABT18794;
XX  DT
XX  16-APR-2003 (first entry)
XX  DE
XX  Aspergillus fumigatus essential gene #1152.
XX  KW
XX  Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX  KW cancer; contamination; biofilm; antibody; immune response; ds.
XX  KW
XX  Aspergillus fumigatus.
XX  OS
XX  WO200286090-A2.
XX  FN
XX  31-OCT-2002.
XX  PD
```

```
23-APR-2002; 2002WO-US013142.
XX  PF
XX  23-APR-2001; 2001US-0285697P.
XX  PR
XX  27-APR-2001; 2001US-0287066P.
XX  PR
XX  05-JUN-2001; 2001US-0295890P.
XX  PR
XX  09-JUL-2001; 2001US-0303899P.
XX  PR
XX  31-AUG-2001; 2001US-0316362P.
XX  PR
XX  (ELIT-) ELITRA PHARM INC.
XX  PA
XX  Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX  PI
XX  WPI; 2003-093124/08.
XX  DR
XX  New purified or isolated nucleic acids of essential genes of Aspergillus
XX  PT fumigatus, useful for treating or preventing infections by A. fumigatus,
XX  PT or for treating a non-infectious disease in a subject e.g. cancer.
XX  PT
XX  Disclosure; Page; 175pp; English.
XX  PS
XX  The invention relates to novel purified or isolated nucleic acids of
XX  CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
XX  CC the invention are used to treat or prevent infections by a pathogenic
XX  CC organism such as A. fumigatus, to treat a non-infectious disease in a
XX  CC subject (e.g. cancer), to prevent or contain contamination of an object
XX  CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
XX  CC biofilm comprising A. fumigatus. The polynucleotides are useful for
XX  CC expressing recombinant protein for characterization, screening or
XX  CC therapeutic use, as markers for host tissues in which the pathogenic
XX  CC organisms invade or reside, for comparing with the DNA sequence of A.
XX  CC fumigatus to identify duplicated genes or paralogues having the same or
XX  CC similar biochemical activity and/or function, for comparing with DNA
XX  CC sequences of other related or distant pathogenic organisms to identify
XX  CC potential orthologous essential or virulence genes, for selecting and
XX  CC making oligomers for attachment to a nucleic acid array for examination
XX  CC of expression patterns, for raising anti-protein antibodies, as an
XX  CC antigen to raise anti-DNA antibodies or to elicit another immune
XX  CC response, and for identifying polynucleotides encoding the other protein
XX  CC with which binding occurs or to identify inhibitors of the binding
XX  CC interaction. The polypeptides may be used to raise antibodies or to
XX  CC elicit immune response, as a reagent in assays designed to quantitatively
XX  CC determine levels of the protein in biological fluids, as a marker for
XX  CC host tissues in which pathogenic organism invade or reside, and to
XX  CC isolate correlative receptors or ligands in the case of virulence
XX  CC factors. This polynucleotide sequence represents one of the essential
XX  CC genes of Aspergillus fumigatus of the invention
XX  CC
XX  SQ Sequence 2944 BP; 676 A; 770 C; 765 G; 733 T; 0 U; 0 Other;
Query Match 1.3%; Score 49; DB 7; Length 2944;
Best Local Similarity 100.0%; Pred. No. 1.5e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3714 AAATGAAAAA 3762
Db 941 AAATGAAAAA 893
RESULT 67
ABT20610/c
ID  ABT20610 standard; DNA; 3362 BP.
XX  AC
XX  ABT20610;
XX  DT
XX  16-APR-2003 (first entry)
XX  DE
XX  Aspergillus fumigatus essential gene #2968.
XX  KW
XX  Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX  KW cancer; contamination; biofilm; antibody; immune response; ds.
XX  KW
XX  Aspergillus fumigatus.
XX  OS
```

XX WO200286090-A2.
 XX 31-OCT-2002.
 XX 23-APR-2002; 2002WO-US013142.
 XX 23-APR-2001; 2001US-0285697P.
 XX 27-APR-2001; 2001US-0287066P.
 XX 05-JUN-2001; 2001US-0295890P.
 XX 09-JUL-2001; 2001US-0303899P.
 XX 31-AUG-2001; 2001US-0316362P.
 XX (ELIT-) ELITRA PHARM INC.
 XX Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
 XX WPI; 2003-093124/08.
 XX New purified or isolated nucleic acids of essential genes of *Aspergillus*
 XX fumigatus, useful for treating or preventing infections by *A. fumigatus*,
 XX or for treating a non-infectious disease in a subject e.g. cancer.
 XX Disclosure; Page; 175pp; English.
 XX The invention relates to novel purified or isolated nucleic acids of
 XX essential genes of *Aspergillus fumigatus*. The isolated nucleic acids of
 XX the invention are used to treat or prevent infections by a pathogenic
 XX organism such as *A. fumigatus*, to treat a non-infectious disease in a
 XX subject (e.g. cancer), to prevent or contain contamination of an object
 XX by *A. fumigatus*, or to prevent or inhibit formation on a surface of a
 XX biofilm comprising *A. fumigatus*. The polynucleotides are useful for
 XX expressing recombinant protein for characterisation, screening or
 XX therapeutic use, as markers for host tissues in which the pathogenic
 XX organisms invade or reside, for comparing with the DNA sequence of *A.*
 XX *fumigatus* to identify duplicated genes or paralogues having the same or
 XX similar biochemical activity and/or function, for comparing with DNA
 XX sequences of other related or distant pathogenic organisms to identify
 XX potential orthologous essential or virulence genes, for selecting and
 XX making oligomers for attachment to a nucleic acid array for examination
 XX of expression patterns, for raising anti-protein antibodies, as an
 XX antigen to raise anti-DNA antibodies or to elicit another immune
 XX response, and for identifying polynucleotides encoding the other protein
 XX with which binding occurs or to identify inhibitors of the binding
 XX interaction. The polypeptides may be used to raise antibodies or to
 XX elicit immune response, as a reagent in assays designed to quantitatively
 XX determine levels of the protein in biological fluids, as a marker for
 XX host tissues in which pathogenic organism invade or reside, and to
 XX isolate correlative receptors or ligands in the case of virulence
 XX factors. This polynucleotide sequence represents one of the essential
 XX genes of *Aspergillus fumigatus* of the invention
 XX Sequence 3362 BP; 774 A; 901 C; 867 G; 820 T; 0 U; 0 Other;
 XX
 XX Query Match 1.3%; Score 49; DB 7; Length 3362;
 XX Best Local Similarity 100.0%; Pred. No. 1.4e-07;
 XX Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 XX
 XX QY 3714 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
 XX |||||
 XX Db 1359 AAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1311
 XX
 XX RESULT 68
 XX ABQ92014
 XX ID ABQ92014 standard; cDNA; 4236 BP.
 XX AC ABQ92014;
 XX XX
 XX DT 04-OCT-2002 (first entry)
 XX DE Human polynucleotide SEQ ID NO 11.
 XX

KW Human; cytostatic; antirheumatic; antiarthritic; vulnary; analgesic;
 KW antiinflammatory; antibacterial; immunosuppressive; antiparkinsonian;
 KW neuroprotective; nootropic; osteopathic; haemostatic; vasotropic;
 KW antitumor; fungicide; antidiabetic; antiaesthetic; antiallergic;
 KW immunostimulant; antiparasitic; secreted protein; transmembrane protein;
 KW cytokine; cell proliferation; cell differentiation; autoimmune disease;
 KW stem cell; growth factor; nervous system disease; neuropathy;
 KW Alzheimer's disease; Parkinson's disease; Huntington's disease;
 KW osteoporosis; severe combined immunodeficiency; SCID; infection;
 KW multiple sclerosis; rheumatoid arthritis; gene therapy; gene; ss.
 XX Homo sapiens.
 XX US2002065394-A1.
 XX 30-MAY-2002.
 XX 22-DEC-2000; 2000US-00745763.
 XX 18-MAR-1998; 98US-00040963.
 XX (JACO/) JACOBS K.
 XX (MCCO/) MCCOY J M.
 XX (LAVA/) LAVALLIE E R.
 XX (COLL/) COLLINS-RACIE L A.
 XX (EVAN/) EVANS C.
 XX (MERE/) MERBERG D.
 XX (TREA/) TREACY M.
 XX (SPAU/) SPAULDING V.
 XX Jacobs K, Mccoy JM, Lavallie ER, Collins-Racie LA, Evans C;
 XX Merberg D, Treacy M, Spaulding V;
 XX WPI; 2002-582343/62.
 XX P-PSDB; ABP61797.
 XX Novel secreted or transmembrane protein and polynucleotide encoding the
 XX protein, useful for diagnosis and treatment of neurological disorders,
 XX cancer, autoimmune diseases, bone disorders and lung or liver fibrosis.
 XX Claim 34; Page 106-108; 284pp; English.
 XX The invention relates to human secreted or transmembrane protein (I),
 XX their fragments and is encoded by specific complementary deoxyribonucleic
 XX acid (cDNA) inserts (II), where the protein is substantially free from
 XX other mammalian proteins. (I) are useful for preventing, treating or
 XX ameliorating a medical condition, especially immunological treatment or
 XX prevention of tumors. (I) exhibits activity relating to angiogenesis,
 XX cytokine, cell proliferation, cell differentiation, antiinflammatory,
 XX stem cell growth factor activity and activin or inhibin-related
 XX activities. (I) can be used to manipulate stem cells in culture to give
 XX rise to neuroepithelial cells that can be used to augment or replace
 XX cells damaged by illness, autoimmune disease, accidental damage or
 XX genetic disorders. (I) induces the proliferation of neural cells and
 XX regeneration of nerve and brain tissue and is useful for the treatment of
 XX central and peripheral nervous system diseases and neuropathies, such as
 XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
 XX lateral sclerosis. (I) is involved in chemotactic or chemokinetic
 XX activity, regulation of haematopoiesis and is useful for treating myeloid
 XX or lymphoid cell disorders, platelet disorders such as thrombocytopenia
 XX and for regeneration of bone, cartilage, tendon, ligament and/or nerve
 XX tissue growth and in tissue repair, healing of burns, incisions, ulcers,
 XX for treating osteoporosis, osteoarthritis, bone degenerative disorders or
 XX periodontal disease. (I) is also useful for gut protection or
 XX regeneration and treatment of lung or liver fibrosis, reperfusion injury
 XX in various tissues, various immune deficiencies and disorders including
 XX severe combined immunodeficiency (SCID), bacterial or fungal infections,
 XX autoimmune disorders e.g. multiple sclerosis, rheumatoid arthritis,
 XX diabetes mellitus, myasthenia gravis, allergic reactions and conditions,
 XX such as asthma or other respiratory problems (II) is useful to express
 XX recombinant protein, as markers for tissues in which the corresponding
 XX protein is preferentially expressed and in gene therapy. The present
 XX sequence is that of a polynucleotide of the invention

XX SQ Sequence 4236 BP; 1330 A; 778 C; 784 G; 1338 T; 0 U; 6 Other;
Query Match 1.3%; Score 49; DB 6; Length 4236;
Best Local Similarity 100.0%; Pred. No. 1.4e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
DB 2421 AAATGAAAAA 2469
RESULT 69
ABT18200/c
ID ABT18200 standard; DNA; 4944 BP.
XX AC ABT18200;
XX DT 16-APR-2003 (first entry)
XX DE Aspergillus fumigatus essential gene #558.
XX KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX KW cancer; contamination; biofilm; antibody; immune response; ds.
XX OS Aspergillus fumigatus.
XX PN WO200286090-A2.
XX PD 31-OCT-2002.
XX PF 23-APR-2002; 2002WO-US013142.
XX PR 23-APR-2001; 2001US-0285697P.
XX PR 27-APR-2001; 2001US-0287066P.
XX PR 05-JUN-2001; 2001US-0295890P.
XX PR 09-JUL-2001; 2001US-0303899P.
XX PR 31-AUG-2001; 2001US-0316362P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
XX PT New purified or isolated nucleic acids of essential genes of Aspergillus
XX PT fumigatus, useful for treating or preventing infections by A. fumigatus,
XX PT or for treating a non-infectious disease in a subject e.g. cancer.
XX PS Disclosure; Page; 175pp; English.
XX CC The invention relates to novel purified or isolated nucleic acids of
XX CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
XX CC the invention are used to treat or prevent infections by a pathogenic
XX CC organism such as A. fumigatus, to treat a non-infectious disease in a
XX CC subject (e.g. cancer), to prevent or contain contamination of an object
XX CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
XX CC biofilm comprising A. fumigatus. The polynucleotides are useful for
XX CC expressing recombinant protein for characterization, screening or
XX CC therapeutic use, as markers for host tissues in which the pathogenic
XX CC organisms invade or reside, for comparing with the DNA sequence of A.
XX CC fumigatus to identify duplicated genes or paralogues having the same or
XX CC similar biochemical activity and/or function, for comparing with DNA
XX CC sequences of other related or distant pathogenic organisms to identify
XX CC potential orthologous essential or virulence genes, for selecting and
XX CC making oligomers for attachment to a nucleic acid array for examination
XX CC of expression patterns, for raising anti-protein antibodies, as an
XX CC antigen to raise anti-DNA antibodies or to elicit another immune
XX CC response, and for identifying polynucleotides encoding the other protein
XX CC determine levels of the protein in biological fluids, as a marker for

CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of Aspergillus fumigatus of the invention
XX SQ Sequence 4944 BP; 1177 A; 1248 C; 1244 G; 1275 T; 0 U; 0 Other;
Query Match 1.3%; Score 49; DB 7; Length 4944;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3714 AAATGAAAAA 3762
DB 1941 AAATGAAAAA 1893
RESULT 70
ABT20014/c
ID ABT20014 standard; DNA; 5362 BP.
XX AC ABT20014;
XX DT 16-APR-2003 (first entry)
XX DE Aspergillus fumigatus essential gene #2372.
XX KW Fungicide; cytostatic; essential gene; Aspergillus fumigatus; infection;
XX KW cancer; contamination; biofilm; antibody; immune response; ds.
XX OS Aspergillus fumigatus.
XX PN WO200286090-A2.
XX PD 31-OCT-2002.
XX PF 23-APR-2002; 2002WO-US013142.
XX PR 23-APR-2001; 2001US-0285697P.
XX PR 27-APR-2001; 2001US-0287066P.
XX PR 05-JUN-2001; 2001US-0295890P.
XX PR 09-JUL-2001; 2001US-0303899P.
XX PR 31-AUG-2001; 2001US-0316362P.
XX PA (ELIT-) ELITRA PHARM INC.
XX PI Jiang B, Tishkoff D, Zamudio C, Eroshkin AM, Hu W, Lemieux SM;
XX WPI; 2003-093124/08.
XX PT New purified or isolated nucleic acids of essential genes of Aspergillus
XX PT fumigatus, useful for treating or preventing infections by A. fumigatus,
XX PT or for treating a non-infectious disease in a subject e.g. cancer.
XX PS Disclosure; Page; 175pp; English.
XX CC The invention relates to novel purified or isolated nucleic acids of
XX CC essential genes of Aspergillus fumigatus. The isolated nucleic acids of
XX CC the invention are used to treat or prevent infections by a pathogenic
XX CC organism such as A. fumigatus, to treat a non-infectious disease in a
XX CC subject (e.g. cancer), to prevent or contain contamination of an object
XX CC by A. fumigatus, or to prevent or inhibit formation on a surface of a
XX CC biofilm comprising A. fumigatus. The polynucleotides are useful for
XX CC expressing recombinant protein for characterization, screening or
XX CC therapeutic use, as markers for host tissues in which the pathogenic
XX CC organisms invade or reside, for comparing with the DNA sequence of A.
XX CC fumigatus to identify duplicated genes or paralogues having the same or
XX CC similar biochemical activity and/or function, for comparing with DNA
XX CC sequences of other related or distant pathogenic organisms to identify
XX CC potential orthologous essential or virulence genes, for selecting and
XX CC making oligomers for attachment to a nucleic acid array for examination
XX CC of expression patterns, for raising anti-protein antibodies, as an
XX CC antigen to raise anti-DNA antibodies or to elicit another immune
XX CC response, and for identifying polynucleotides encoding the other protein
XX CC interaction. The polypeptides may be used to raise antibodies or to
XX CC elicit immune response, as a reagent in assays designed to quantitatively
XX CC determine levels of the protein in biological fluids, as a marker for

CC with which binding occurs or to identify inhibitors of the binding
CC interaction. The polypeptides may be used to raise antibodies or to
CC elicit immune response, as a reagent in assays designed to quantitatively
CC determine levels of the protein in biological fluids, as a marker for
CC host tissues in which pathogenic organism invade or reside, and to
CC isolate correlative receptors or ligands in the case of virulence
CC factors. This polynucleotide sequence represents one of the essential
CC genes of *Aspergillus fumigatus* of the invention
XX
SQ Sequence 5362 BP; 1321 A; 1343 C; 1334 G; 1364 T; 0 U; 0 Other;

Query Match 1.3%; Score 49; DB 7; Length 5362;
Best Local Similarity 100.0%; Pred. No. 1.3e-07;
Matches 49; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3714 AAATGAAAAA...AAA 3762
DB 2359 AAATGAAAAA...AAA 2311

RESULT 71
AAH69364
ID AAH69364 standard; cDNA; 295 BP.
XX
AC AAH69364;
XX
DT 19-SEP-2001 (first entry)
XX
DE Human cervical cancer marker nucleic acid 638.
XX
KW Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS Homo sapiens.
XX
PN WO200142467-A2.
XX
PD 14-JUN-2001.
XX
PF 08-DEC-2000; 2000WO-US033312.
XX
PR 08-DEC-1999; 99US-0169681P.
PR 21-DEC-1999; 99US-0171350P.
PR 14-MAR-2000; 2000US-0189315P.
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI; 2001-375006/39.
XX

PT New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.

XX
PS Claim 1; Page 215; 1051pp; English.
XX
CC The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX

XX
SQ Sequence 295 BP; 133 A; 46 C; 47 G; 69 T; 0 U; 0 Other;
Query Match 1.3%; Score 48; DB 4; Length 295;
Best Local Similarity 100.0%; Pred. No. 5.1e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA...AAA 3762
DB 233 AATGAAAAA...AAA 280
RESULT 72
ABX40871
ID ABX40871 standard; cDNA; 337 BP.
XX
AC ABX40871;
XX
DT 20-FEB-2003 (first entry)
XX
DE Bovine EST associated with lactation/muscle/fat deposition #6036.
XX
KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
KW muscle deposition; fat deposition; genome mapping; gene identification;
KW gene analysis; cattle breeding.
XX
OS Bos Taurus.
XX
PN US2002137139-A1.
XX
PD 26-SEP-2002.
XX
PF 24-SEP-2001; 2001US-00960352.
XX
PR 12-JAN-1999; 99US-0115707P.
PR 11-JAN-2000; 2000US-00480902.
XX
PA (BYAT/) BYATT J C.
PA (MATH/) MATHIALAGAN N.
PA (TAON/) TAO N.
PA (WARE/) WARREN W C.
XX
PI Byatt JC, Mathialagan N, Tao N, Warren WC;
XX
DR WPI; 2003-110599/10.
XX

PT New nucleic acid associated with lactation, and muscle and fat
PT deposition, useful for genome mapping, gene identification and analysis,
PT cattle breeding, or for genetically improving cattle.

PS Claim 2; SEQ ID NO 6036; 245pp; English.

XX
CC The invention relates to a purified nucleic acid molecule associated with
CC lactation or muscle and fat deposition (designated LMFD), derived from
CC cattle, and the LMFD nucleic acid can specifically hybridize to a second
CC nucleic acid molecule comprising any of 15112 nucleotide sequences,
CC appearing as ABX34836-ABX49947, or complements of them. Also included are
CC ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
CC acid linked to a promoter and a 3' non-translated sequence that
CC functions in the cell to cause termination of transcription and addition
CC of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
CC (2) determining a level or pattern of a molecule in a bovine cell or
CC tissue comprising: (a) incubating a marker nucleic acid (comprising any
CC of the 15112 nucleic acid sequences or its complement or fragment) with a
CC complementary nucleic acid molecule obtained from the bovine cell or
CC tissue, where hybridisation between the marker nucleic acid and the
CC complementary nucleic acid permits the detection of the molecule; and (b)
CC detecting the level or pattern of the complementary nucleic acid, where
CC the detection of the complementary nucleic acid is predictive of the
CC level or pattern of the molecule. The LMFD nucleic acid is used for
CC determining a level or pattern of a molecule in a bovine cell or tissue.
CC It is useful for genome mapping, gene identification and analysis, cattle
CC breeding, preparation of constructs for use in cattle gene expression, or
CC for genetically improving cattle. The present sequence is one of the
CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
CC present sequence was not shown in the specification but was obtained in
CC electronic format from the USPTO web site:
CC seqdata.uspto.gov/sequence.html?DocID=20020137139

XX
SQ Sequence 337 BP; 150 A; 35 C; 40 G; 112 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 7; Length 337;
Best Local Similarity 100.0%; Pred. No. 5e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
DB 138 AATGAAAAA
3762
185

RESULT 73
ACH29107
ID ACH29107 standard; cDNA; 401 BP.

AC ACH29107;

XX 13-OCT-2003 (first entry)

XX Human adult spleen cDNA #126.

XX Human; ss; sequencing by hybridisation; SBH; expressed sequence tag; EST;
KW genome mapping; biodiversity; Genetic disorder.

XX Homo sapiens.

XX US2003073623-A1.

XX 17-APR-2003.

XX 30-JUL-2001; 2001US-00918995.

XX 30-JUL-2001; 2001US-00918995.

XX (DRMA/) DRMANAC R T.

XX (LABA/) LABAT I.

XX (STAC/) STACHE-CRAIN B.

XX (DICK/) DICKSON M C.

XX (JONE/) JONES L W.

XX Drmanac RT, Labat I, Stache-Crain B, Dickson MC, Jones LW;

XX WPI; 2003-615964/58.

XX New polynucleotide sequences obtained from various cDNA libraries, useful
PT as hybridization probes, as oligomers for PCR, for chromosome and gene
PT mapping, in the recombinant production of protein, or in generating
PT antisense DNA or RNA.

XX Claim 1; SEQ ID NO 16319; 44pp; English.

XX The invention relates to an isolated polynucleotide comprising any one of
CC 38043 cDNA sequences, appearing as ACH12789-ACH50831, whose sequence was
CC determined by the technique of SBH (sequencing by hybridisation). Also
CC included is a purified polypeptide comprising a sequence corresponding to
CC a reading frame of the novel polynucleotide. The nucleic acid sequences
CC are useful in diagnostics as expressed sequence tags (EST) for
CC identifying expressed genes or for physical mapping of the human genome,
CC in forensics, in assessing biodiversity, or in identifying mutations
CC responsible for genetic disorders and other traits. The nucleotide
CC sequences are also useful as hybridisation probes, as oligomers for PCR,
CC for chromosome and gene mapping, in the recombinant production of
CC protein, or in generating antisense DNA or RNA. The purified polypeptide
CC is useful for generating antibodies specific for it. The present sequence
CC is one of the 38043 isolated cDNA/EST sequences. Note: The sequence data
CC for this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from USPTO at
CC seqdata.uspto.gov/sequence.html?DocID=20030073623

XX Sequence 401 BP; 243 A; 47 C; 50 G; 61 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 401;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
DB 84 AATGAAAAA
3762
131

RESULT 74
ABV44841/C
ID ABV44841 standard; cDNA; 404 BP.

XX AC ABV44841;

XX 16-SEP-2002 (first entry)

XX Human prostate expression marker cDNA 44832.

XX Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.

XX Homo sapiens.

XX WO200160860-A2.

XX 23-AUG-2001.

XX 20-FEB-2001; 2001WO-US005171.

XX 17-FEB-2000; 2000US-0183319P.

XX 16-MAR-2000; 2000US-0189862P.

XX 25-MAY-2000; 2000US-0207454P.

XX 09-JUN-2000; 2000US-0211314P.

XX 18-JUL-2000; 2000US-0219007P.

XX 13-DEC-2000; 2000US-0255281P.

XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

XX Schlegel R, Endege WO, Monahan JE;

XX WPI; 2001-662795/76.

XX Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.

XX Claim 1; Page 8889; 11750pp; English.

XX The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker

XX Sequence 404 BP; 119 A; 77 C; 67 G; 141 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 5; Length 404;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
DB 98 AATGAAAAA
3762
51

RESULT 75

AA187343
ID AA187343 standard; cDNA; 418 BP.

```
XX AC AAT87343;
XX OS
XX DT 06-NOV-2001 (first entry)
XX PN Human polynucleotide SEQ ID NO 7403.
XX DE
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
XX KW nervous system disorders; arthritis; inflammation; ss.
XX OS Homo sapiens.
XX FN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX PT P-PSDB; AAO07412.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 1; SEQ ID NO 7403; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and/or
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 418 BP; 222 A; 43 C; 62 G; 91 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 4; Length 418;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 272 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 319

RESULT 76
AAI89543
ID AAI88543 standard; cDNA; 419 BP.
XX AC AAI88543;
XX DT 06-NOV-2001 (first entry)
XX PN Human polynucleotide SEQ ID NO 8603.
XX DE
XX KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
XX KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
XX KW tissue growth factor; immunomodulatory; cancer; leukaemia;
```

```
KW nervous system disorders; arthritis; inflammation; ss.
XX OS Homo sapiens.
XX PN WO200164835-A2.
XX PD 07-SEP-2001.
XX PF 26-FEB-2001; 2001WO-US004927.
XX PR 28-FEB-2000; 2000US-00515126.
XX PR 18-MAY-2000; 2000US-00577409.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang YT, Liu C, Drmanac RT;
XX DR WPI; 2001-514838/56.
XX PT P-PSDB; AAO08612.
XX PT Isolated nucleic acids and polypeptides, useful for preventing diagnosing
XX PT and treating e.g. leukemia, inflammation and immune disorders.
XX PS Claim 1; SEQ ID NO 8603; 1399pp + Sequence Listing; English.
XX CC The invention relates to human polynucleotides (AAI79941-AAI93841) and
XX CC the encoded proteins (AAO0010-AAO13910) that exhibit activity relating to
XX CC cytokine, cell proliferation or cell differentiation or which may induce
XX CC production of other cytokines in other cell populations. The
XX CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
XX CC peptide therapy. The polypeptides have various cytokine-like activities,
XX CC e.g. stem cell growth factor activity, haematopoiesis regulating
XX CC activity, tissue growth factor activity, immunomodulatory activity and/or
XX CC activin/inhibin activity and may be useful in the diagnosis and/or
XX CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
XX CC inflammation. Note: The sequence data for this patent did not form part
XX CC of the printed specification, but was obtained in electronic format
XX CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX SQ Sequence 419 BP; 174 A; 78 C; 96 G; 81 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 4; Length 419;
Best Local Similarity 100.0%; Pred. No. 4.8e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 219 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 266

RESULT 77
AAH71167
ID AAH71167 standard; cDNA; 546 BP.
XX AC AAH71167;
XX DT 19-SEP-2001 (first entry)
XX PN Human cervical cancer marker nucleic acid 2441.
XX DE Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX KW Homo sapiens.
XX OS Homo sapiens.
XX PN WO200142467-A2.
XX PD 14-JUN-2001.
XX PF 08-DEC-2000; 2000WO-US033312.
XX PR 08-DEC-1999; 99US-0169681P.
XX PR 21-DEC-1999; 99US-0171350P.
XX PR 14-MAR-2000; 2000US-0189315P.
```



```
PR 12-MAY-2000; 2000US-0203791P.
PR 09-JUN-2000; 2000US-0210600P.
PR 21-JUL-2000; 2000US-0220114P.
XX (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
XX WPI; 2001-375006/39.
XX
XX New isolated nucleic acid for diagnosing and treating cervical cancer and
PT for assessing and detecting compounds for treating the cancer.
XX
XX Claim 1; Page 512; 1051pp; English.
XX
XX The invention relates to novel genes (AAH68727-AAH73383) associated with
CC cervical cancer with cytostatic activity. The nucleic acids and encoded
CC polypeptides are useful: to assess if a patient is afflicted with
CC cervical cancer or has a pre-malignant condition; to monitor the
CC progression of cervical cancer or a premalignant condition in a patient;
CC and to select and/or assess the efficacy of a compound or therapy for
CC inhibiting cervical cancer in a patient. The nucleic acids may also be
CC useful for gene therapy
XX
XX Sequence 546 BP; 240 A; 92 C; 83 G; 130 T; 0 U; 1 Other;
XX
Query Match 1.3%; Score 48; DB 4; Length 546;
Best Local Similarity 100.0%; Pred. No. 4.5e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAA 3762
DB 248 AATGAAAAA 295
RESULT 78
AAA26393
ID AAA26393 standard; cDNA; 619 BP.
XX
XX AAA26393;
XX
DT 29-JUN-2000 (first entry)
XX
DE Human secreted protein gene 48 SEQ ID NO:58.
XX
XX Human; secreted protein; diagnosis; cytostatic; immunosuppressive;
XX antiHIV; antiinflammatory; neutropic; neuroprotective; antiallergic;
XX osteopathic; antiarthritic; antibacterial; antidiabetic; antiasthma;
XX antiporiatic; cardiant; gene therapy; cancer; neurological disorder;
XX immune disease; inflammation; blood disorder; tumour; ss.
XX
XX Homo sapiens.
XX
XX WO200006698-A1.
XX
XX 10-FEB-2000.
XX
XX 29-JUL-1999; 99WO-US017130.
XX
XX 30-JUL-1998; 98US-0094657P.
XX
XX 05-AUG-1998; 98US-0095486P.
XX
XX 06-AUG-1998; 98US-0095454P.
XX
XX 06-AUG-1998; 98US-0095455P.
XX
XX 12-AUG-1998; 98US-0096319P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Konatsoulis GA, Rosen CA, Ruben SM, Duan R, Moore PA, Shi Y;
XX Lafleur D, Wei Y, Ni J, Florence KA, Young PE, Brewer LA;
XX Soppet DR, Endress GA, Ebner R, Olsen HS, Mucenski N;
XX
XX WPI; 2000-195282/17.
XX
XX P-PSDB; AAY91498.
XX
XX New isolated human genes and the secreted polypeptides they encode,
PT useful for diagnosis and treatment of e.g. cancers, neurological
PT disorders, immune diseases, inflammation or blood disorders.
XX
XX Claim 1; Page 405-406; 634pp; English.
XX
XX The polynucleotide sequences given in AAA26346 to AAA26458 encode the
CC human secreted proteins given in AAY91451 to AAY91691. The human secreted
CC proteins can have activities based on the tissues and cells they are
CC expressed in. Examples of the activities are: cytostatic;
CC immunosuppressive; antiHIV; antiinflammatory; neutropic; neuroprotective;
CC antiallergic; osteopathic; antiarthritic; antibacterial; antidiabetic;
CC antiasthma; antiporiatic; and cardiant. The polynucleotides and their
CC corresponding secreted proteins are useful for preventing, treating or
CC ameliorating medical conditions, e.g. by protein or gene therapy. Also
CC pathological conditions can be diagnosed by determining the amount of the
CC proteins in a sample or by determining the presence of mutations in the
CC polynucleotides. Specific uses are described for each of the
CC polynucleotides, based on which tissues they are most highly expressed
CC in, and include developing products for the diagnosis or treatment of
CC cancer, tumours, neurodegenerative disorders, developmental abnormalities
CC and foetal deficiencies, blood disorders, diseases of the immune system,
CC autoimmune diseases, hepatic and renal disease, inflammation, allergies,
CC Alzheimer's and behavioural disorders, schizophrenia, osteoporosis,
CC arthritis, infections, AIDS, spinal cord injuries, transplant rejection,
CC diabetes, asthma, sepsis, acne, psoriasis, cardiovascular disorders,
CC reproductive disorders, gastrointestinal disorders, respiratory disorders
CC and metabolic disorders. The proteins or polynucleotides can also be used
CC as food additives or preservatives. The proteins are also useful for
CC identifying their binding partners. AAA26337 to AAA26345 and AAY91450 are
CC sequences used in the exemplification of the present invention
XX
XX Sequence 619 BP; 162 A; 168 C; 110 G; 177 T; 0 U; 2 Other;
XX
Query Match 1.3%; Score 48; DB 3; Length 619;
Best Local Similarity 100.0%; Pred. No. 4.4e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAA 3762
DB 566 AATGAAAAA 613
RESULT 79
ABN98811/c
ID ABN98811 standard; DNA; 865 BP.
XX
XX ABN98811;
XX
XX 01-AUG-2002 (first entry)
XX
XX Arabidopsis thaliana expressed polynucleotide SEQ ID NO 579.
XX
XX Arabidopsis thaliana; plant; insecticide; fungicide; transgenic; stress;
XX disease; crop; thale cress; tolerance factor; insect; pathogen;
XX nutrition; ds.
XX
XX Arabidopsis thaliana.
XX
XX US2002023281-A1.
XX
XX 21-FEB-2002.
XX
XX 26-JAN-2001; 2001US-00770445.
XX
XX 27-JAN-2000; 2000US-0178472P.
XX
XX (GORL/) GORLACH J.
XX (ANY/) AN Y.
XX (HAMI/) HAMILTON C M.
XX (PRIC/) PRICE J L.
XX (RAIN/) RAINES T M.
```



```
DE Human secreted protein coding sequence SEQ ID NO: 33.
XX
XX Human; secreted protein; cytokine; cell proliferation;
KW nutritional supplement; immune modulation; autoimmune disorder;
KW haematopoiesis regulation; tissue growth; haemostasis; inflammation; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 93..266
XX /tag= a
XX /product= "secreted protein"
XX sig_peptide 135..173
XX /tag= b
XX mat_peptide 174..263
XX /tag= c
XX
XX WO200049134-A1.
XX
XX 24-AUG-2000.
XX
XX 18-FEB-2000; 2000WO-US004340.
XX
XX 19-FEB-1999; 99US-0120680P.
XX 23-APR-1999; 99US-00398733.
XX 17-AUG-1999; 99US-0149639P.
XX 23-SEP-1999; 99US-0155686P.
XX 01-OCT-1999; 99US-0157247P.
XX 29-NOV-1999; 99US-0167822P.
XX 29-NOV-1999; 99US-0167823P.
XX 15-FEB-2000; 2000US-0182711P.
XX
XX (ALPH-) ALPHAGENE INC.
XX
XX PI Valenzuela D, Yuan O, Hoffman H, Hall J, Rapijko P;
XX WPI; 2000-549267/50.
XX P-PSDB; AAB23617.
XX
XX New secreted proteins and polynucleotides encoding them, which are
XX derived from Homosapiens, useful for therapy, diagnosis, and research, as
XX well as nutritional sources or supplements.
XX
XX Claim 42; Page 264; 309pp; English.
XX
XX The present sequence is the coding sequence for a human secreted protein.
XX The sequence was isolated from an adult lung cDNA library. The proteins
XX and coding sequences of the invention can be used in the isolation of
XX similar genes and proteins, in the elucidation of their function in vivo,
XX and to treat a number of conditions. It is possible that they may have
XX uses as nutritional supplements, as cytokine or cell proliferation
XX factors, in immune modulation, where they may be used to treat immune and
XX autoimmune diseases, as haematopoiesis regulators (treating myeloid or
XX lymphoid cell deficiencies), in the promotion of tissue growth, they may
XX have chemokine or chemotactic activity, haemostatic or thrombolytic
XX activity, or anti-inflammatory activity
XX
XX Sequence 1136 BP; 351 A; 253 C; 249 G; 283 T; 0 U; 0 Other;
SQ
Query Match 1.3%; Score 48; DB 3; Length 1136;
Best Local Similarity 100.0%; Pred. No. 3.9e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1080 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1127
RESULT 82
AAC79042
ID AAC79042 standard; DNA; 1467 BP.
XX
XX AAC79042;
XX
XX 14-FEB-2001 (first entry)
XX
XX Human secreted protein gene 46 clone HPWA089.
XX
XX Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
XX anti-allergic; hepatotropic; antidiabetic; anti-inflammatory; anti-ulcer;
XX vulnerrary; anticonvulsant; antibacterial; antifungal; antiparasitic;
XX cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
XX neurological disease; infection; human; secreted protein; ss.
XX
XX Homo sapiens.
XX
XX WO200058358-A1.
XX
XX 05-OCT-2000.
XX
XX 23-MAR-2000; 2000WO-US007725.
XX
XX 26-MAR-1999; 99US-0126602P.
XX 14-JAN-2000; 2000US-0176063P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Rosen CA, Ruben SM, Komatsoulis G;
XX WPI; 2000-594640/56.
XX P-PSDB; AAB44379.
XX
XX Fourty nine nucleic acid molecules encoding human secreted proteins,
XX useful in the prevention, treatment and diagnosis of cancer, immune
XX disorders, cardiovascular disorders and neurological diseases.
XX
XX Claim 1; Page 336-337; 367pp; English.
XX
XX The invention relates to the isolation of genes AAC78997-C79045 encoding
XX 49 human secreted proteins AAB4335-B44382. The genes can be used to
XX generate fusion proteins by linking to the gene for the human
XX immunoglobulin G Fc portion (AAC78988) for increasing the stability of
XX the fusion protein as compared to the human protein only. The genes and
XX conditions are useful for preventing, ameliorating or treating medical
XX conditions, e.g. by protein or gene therapy. The genes are isolated from
XX a range of human tissues disclosed in the specification. The nucleic
XX acids, proteins, antibodies and (ant)agonists are useful in the
XX diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer, and other cancers of the adrenal gland, bone, bone
XX marrow, breast, gastrointestinal tract, liver, lung, or urogenital; (b)
XX immune disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections
XX
XX Sequence 1467 BP; 429 A; 288 C; 294 G; 447 T; 0 U; 9 Other;
SQ
Query Match 1.3%; Score 48; DB 3; Length 1467;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1414 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1461
RESULT 83
ABK87557
ID ABK87557 standard; cDNA; 1510 BP.
XX
XX ABK87557;
XX
XX 24-SEP-2002 (first entry)
XX
```

```

XX DE cDNA encoding matrix metalloproteinase inhibitor, N-Tes.
XX KW Human; N-Tes; membrane-type matrix metalloproteinase inhibitor; cancer;
XX KW cancer infiltration; angiogenesis; Alzheimer's disease; arthritis;
XX KW glioma; cytostatic; antiarthritic; nootropic; neuroprotective;
XX KW angiogenic; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 108..1049
XX FT /*tag= a
XX FT /product= "N-Tes"
XX FT /note= "Matrix metalloproteinase inhibitor"
XX PN WO200257448-A1.
XX PD 25-JUL-2002.
XX PF 27-DEC-2001; 2001WO-JP011529.
XX PR 27-DEC-2000; 2000JP-00398817.
XX PA (DAII-) DAIICHI FINE CHEM CO LTD.
XX PI Sato H, Acki T;
XX DR WPI; 2002-538473/57.
XX DR P-PSDB; AAU98506.
XX PT New human N-Tes polypeptide and its partial sequences inhibiting matrix
XX PT metalloproteinase for treatment and diagnosis of glioma and other
XX PT neurological and neoplastic disorders.
XX PS Claim 7; Page 114-115; 129pp; Japanese.
XX CC The invention describes polypeptides and their salts containing the
XX CC sequence of human N-Tes polypeptide or derived from it, and inhibiting
XX CC the activity of membrane-type matrix metalloproteinase. The polypeptides
XX CC are used for the treatment, prevention and diagnosis of cancer and cancer
XX CC infiltration, angiogenesis, Alzheimer's disease and arthritis, and
XX CC especially of glioma. This sequence encodes the novel N-Tes polypeptide
XX CC described in the invention
XX SQ Sequence 1510 BP; 507 A; 291 C; 335 G; 377 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 6; Length 1510;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1455 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1502

RESULT 84
AAA46483
ID AAA46483 standard; cDNA; 1544 BP.
XX AC AAA46483;
XX DT 04-SEP-2000 (first entry)
XX DE cDNA encoding a filament-like protein 4.
XX KW MPPI binding factor 1; MAF1; transgenic plant; chimeric gene; NMPI;
XX KW nuclear matrix protein-1; FLIP; filament-like protein;
XX KW transcriptional activator; protein composition; crop growth;
XX KW crop protection; ss.
XX OS Lycopersicon esculentum.

```

```

FH Key Location/Qualifiers
FT CDS 2..1021
FT /*tag= a
FT /product= "filament-like protein 4"
PN WO200028054-A2.
XX PD 18-MAY-2000.
XX PF 04-NOV-1999; 99WO-US025993.
XX PR 06-NOV-1998; 98US-00187999.
XX PA (DUPO) DU PONT DE NEMOURS & CO S I.
XX PI Gindullis F, Meier I;
XX DR WPI; 2000-376565/32.
XX DR P-PSDB; AAY93406.
XX PT Regulating gene expression in plant cells, useful e.g. for expressing new
XX PT traits, by introducing genes for transcriptional activators that interact
XX PT with nuclear matrix proteins.
XX PS Claim 24; Page 72; 82pp; English.
XX CC The present sequence encodes a filament-like protein 4 (FLIP4), which is
XX CC used in the course of the invention. The specification describes a method
XX CC whereby gene expression in a stably transformed transgenic plant cell is
XX CC regulated using a chimeric gene that encodes NMPI (nuclear matrix protein
XX CC -1) or FLIP. Two chimeric genes are combined into the genome, the first
XX CC gene comprising a promoter linked to a DNA binding domain, a coding
XX CC sequence or its complement, and a polyadenylation sequence. The second
XX CC gene comprises a promoter, a DNA binding domain, NMPI or FLIP
XX CC polynucleotides, and a polyadenylation sequence. Expression of the first
XX CC gene regulates expression of the second. NMPI, FLIP and related proteins
XX CC are parts of the nuclear skeleton and function as transcriptional
XX CC activators. The methods are used to regulate gene expression in plants,
XX CC e.g. to express novel traits in transgenic plants that may result in new
XX CC products useful in foods, pharmaceuticals and materials, or for
XX CC suppressing endogenous genes to alter the protein composition, and to
XX CC derive new phenotypes beneficial for crop growth and development. NMPI,
XX CC FLIP and related proteins are used to alter level of expression of MPPI,
XX CC binding factor 1 (MAFI) or MAF1-binding proteins in plants by co-
XX CC suppression or overexpression, or to screen for compounds that inhibit
XX CC activity of MPPI- or MAF1-binding proteins, which are potentially useful
XX CC as crop protection agents. They can also be used to study the plant
XX CC nucleus matrix
XX SQ Sequence 1544 BP; 492 A; 259 C; 367 G; 426 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 3; Length 1544;
Best Local Similarity 100.0%; Pred. No. 3.7e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
DB 1470 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1517

RESULT 85
ABK33555
ID ABK33555 standard; cDNA; 1837 BP.
XX AC ABK33555;
XX DT 08-MAY-2002 (first entry)
XX DE cDNA encoding human PRO protein, Seq ID No 39.
XX KW Human; secreted protein; PRO; tumour; lung cancer; colon cancer;
XX KW breast cancer; prostate tumour; rectal tumour; liver tumour;
XX KW pericyte cell proliferation; chondrocyte cell proliferation;

```

tumour necrosis factor-alpha; gene; ss.

Homo sapiens.

WO200208288-A2.

31-JAN-2002.

29-JUN-2001; 2001WO-US021066.

20-JUL-2000; 2000US-0219556P.

25-JUL-2000; 2000US-0220585P.

25-JUL-2000; 2000US-0220605P.

25-JUL-2000; 2000US-0220607P.

25-JUL-2000; 2000US-0220624P.

25-JUL-2000; 2000US-0220638P.

25-JUL-2000; 2000US-0220654P.

25-JUL-2000; 2000US-0220656P.

26-JUL-2000; 2000US-0220893P.

28-JUL-2000; 2000WO-US020710.

01-AUG-2000; 2000US-0222425P.

22-AUG-2000; 2000US-0227133P.

23-AUG-2000; 2000WO-US023522.

24-AUG-2000; 2000WO-US023328.

10-NOV-2000; 2000WO-US030873.

28-NOV-2000; 2000US-0253646P.

01-DEC-2000; 2000WO-US032678.

20-DEC-2000; 2000US-00747259.

20-DEC-2000; 2000WO-US034856.

28-FEB-2001; 2001WO-US006520.

01-MAR-2001; 2001WO-US006666.

22-MAR-2001; 2001US-00816744.

10-MAY-2001; 2001US-00854208.

10-MAY-2001; 2001US-00854280.

25-MAY-2001; 2001WO-US017092.

(GETH) GENENTECH INC.

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ,

Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

WPI; 2002-172001/22.

P-PSDB; AAU83611.

One hundred and twenty two nucleic acids encoding PRO polypeptides,

useful for treating a PRO related disorder and for diagnosing tumors such

as lung cancer, colon cancer, breast tumor, prostate tumor, rectal tumor

or liver tumor.

Claim 2; Fig 39; 359pp; English.

The invention relates to one hundred and twenty two nucleic acids

encoding PRO polypeptides. The sequences of the 122 PRO polynucleotides

encode human secreted proteins. The PRO nucleic acids, polypeptides,

agonists and antagonists are useful for treating a PRO related disorder.

The PRO polypeptides are useful for diagnosing tumors, especially lung

cancer, colon cancer, breast tumor, prostate tumor, rectal tumor or

liver tumor. The PRO polypeptides are useful for stimulating the

proliferation of, or gene expression, in pericyte cells, for stimulating

the proliferation or differentiation of chondrocyte cells, for

stimulating the release of tumour necrosis factor-alpha from human blood,

for stimulating or inhibiting the proliferation of normal human dermal

fibroblast cells. The PRO polypeptide may also be used as molecular

weight markers and for tissue typing. The PRO nucleic acids have

applications in molecular biology, including use as hybridisation probes,

and in chromosome and gene mapping. ABK33536-ABK33657 represent human PRO

protein coding sequences of the invention

Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 6; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

3715 AATGAAAAA

DB

1729 AATGAAAAA

RESULT 86

ACA66860

ID

ACA66860 standard; cDNA; 1837 BP.

XX

AC

ACA66860;

XX

23-JUN-2003

(first entry)

XX

DT

cDNA encoding human PRO polypeptide #20.

XX

DE

Human; PRO polypeptide; secreted and transmembrane protein;

XX

KW

anti-PRO antibody; diagnostic assay; gene expression; tumour; cytostatic;

XX

Gene; ss.

XX

OS

Homo sapiens.

XX

PN

US2003036635-A1.

XX

PD

20-FEB-2003.

XX

XX

28-AUG-2002; 2002US-00230163.

XX

PF

25-JUL-2000; 2000US-0220638P.

XX

PR

01-JUN-2001; 2001WO-US017800.

XX

PR

23-JUN-2001; 2001WO-US021066.

XX

PR

03-APR-2002; 2002US-00119480.

XX

XX

(GETH) GENENTECH INC.

XX

Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX

PI

Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX

DR

WPI; 2003-342045/32.

XX

DR

P-PSDB; ABU80758.

XX

PT

One hundred and twenty two nucleic acids encoding PRO polypeptides,

XX

PT

useful for the manufacture of a medicament for diagnosing or treating

XX

PT

tumor.

XX

PS

Claim 2; Fig 39; 314pp; English.

XX

CC

The present invention relates to the isolation of novel human PRO

XX

CC

polypeptides, and the polynucleotide sequences encoding them. The PRO

XX

CC

polypeptides are secreted and transmembrane proteins. The PRO

polypeptides and polynucleotides are useful for preparing a medicament

XX

CC

useful in the diagnosis and treatment of tumours. Anti-PRO antibodies are

XX

CC

useful in diagnostic assays for PRO, by detecting its expression in

XX

CC

specific cells, tissues or serum, and for affinity purification of PRO

XX

CC

from recombinant cell culture or natural sources. ACA66841-ACA66962

XX

CC

represent cDNA sequences encoding the human PRO polypeptides of the

XX

CC

invention. Note: The sequence data for this patent was obtained in

XX

CC

electronic format directly from the USPTO web site at

XX

CC

segdata.uspto.gov/psipsDiEntry.html

XX

XX

Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

XX

Query Match 1.3%; Score 48; DB 7; Length 1837;

XX

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

XX

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY

3715 AATGAAAAA

DB

1729 AATGAAAAA

RESULT 87

ACD68612

```

ID XX ACD68612 standard; cDNA; 1837 BP.
AC XX ACD68612;
XX
DT 17-SEP-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO10111 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; cytotatic;
KW antiarthritic; osteopathic; gene therapy; TNF-Agonist-Alpha;
KW chondrocyte stimulator; pericyte stimulator; fibroblast modulator;
KW pharmaceutical; diagnostic; biosensor; bioreactor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; bone disorder; cartilage disorder; sports injury;
KW arthritis; wound; gene; ss.
XX
OS Homo sapiens.
XX
PN US2003045687-A1.
XX
PD 06-MAR-2003.
XX
PF 12-AUG-2002; 2002US-00218631.
XX
PR 01-JUN-2001; 2001WO-US017800.
PR 29-JUN-2001; 2001WO-US021086.
PR 09-APR-2002; 2002US-00119480.
XX
PA (GETH ) GENENTECH INC.
XX
PI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX WPI; 2003-512315/48.
DR P-PSDB; ABO33724.
DR
XX
XX New genes, and its encoded secreted and transmembrane polypeptides,
PT useful for stimulating Tumor Necrosis Factor alpha, or chondrocyte or
PT pericyte proliferation, especially for treating lung tumors, arthritis or
PT wounds in a mammal.
XX
XX Claim 2; Fig 39; 314pp; English.
XX
XX The invention describes an isolated nucleic acid molecule comprising a
CC sequence with at least 80% identity to: (a) a nucleotide encoding any of
CC 122 PRO (secreted and transmembrane) polypeptides whose sequences are
CC fully defined in the specification; or (b) any of 122 nucleotide
CC sequences having e.g. 4834, 2504 or 1759 bp fully defined in the
CC specification; or the full length coding sequence of any these 122
CC nucleotide sequences. The PRO polypeptides or polynucleotides are useful
CC as pharmaceuticals, diagnostics, biosensors or bioreactors. These are
CC particularly useful for detecting tumours (e.g. lung tumour, colon
CC tumour, breast tumour, prostate tumour, rectal tumour, or liver tumour)
CC in a mammal, for stimulating the release of TNF-alpha from human blood,
CC for stimulating the proliferation or differentiation of chondrocyte
CC cells, for stimulating proliferation of pericyte cells, or for modulating
CC normal human dermal fibroblast proliferation. The PRO nucleic acid or
CC polypeptide is also useful for treating tumours or various bone and/or
CC cartilage disorders (e.g. sports injuries or arthritis), or wounds. The
CC PRO polypeptides are useful in drug screening, particularly as targets
CC for therapeutic intervention in these diseases, and in the diagnostic
CC determination of the presence of these diseases. The PRO polypeptides are
CC also useful as molecular weight markers, or for chromosome
CC screening libraries of human cDNA, genomic DNA or mRNA. The PRO genes may
CC also be used in gene therapy, particularly for replacing a defective
CC gene. This sequence encodes a novel human secreted and transmembrane PRO
CC polypeptide
XX
XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;
SQ Query Match 1.3%; Score 48; DB 7; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 RAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
XX |||||
DB 1729 RAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
XX |||||

RESULT 88
ACA68516
ID ACA68516 standard; cDNA; 1837 BP.
XX
XX ACA68516;
AC
XX
DT 25-JUN-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO10111 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; cardiant; cytotatic;
KW antiangiogenic; hypotensive; vulnerry; antiarteriosclerotic;
KW gene therapy; cardiovascular disorder; endothelial disorder;
KW angiogenic disorder; cardiac hypertrophy; trauma; cancer;
KW age-related macular degeneration; atherosclerosis; hypertension;
KW arterial restenosis; rheumatoid arthritis; angina; myocardial infarction;
KW thrombophlebitis; lymphangitis; tumour angiogenesis; breast carcinoma;
KW liver carcinoma; wound healing; chromosome mapping; gene mapping; gene;
XX ss.
XX Homo sapiens.
XX
XX US2003088063-A1.
XX
XX 08-MAY-2003.
XX
XX 12-AUG-2002; 2002US-00219003.
XX
XX 25-JUL-2000; 2000US-0220664P.
PR 01-JUN-2001; 2001WO-US017800.
PR 23-JUN-2001; 2001WO-US021066.
PR 09-APR-2002; 2002US-00119480.
XX
XX (GETH ) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX
XX WPI; 2003-393229/37.
DR P-PSDB; ABUS2057.
XX
XX One hundred and eighty seven nucleic acids encoding PRO polypeptides,
PT useful in diagnosis and treatment of cardiovascular (e.g. myocardial
PT infarction), endothelial or angiogenic disorders in a mammal.
XX
XX Claim 2; Fig 39; 314pp; English.
XX
XX The invention describes one hundred and eighty seven nucleic acids
CC encoding novel human secreted and transmembrane (PRO) polypeptides. The
CC PRO nucleic acids, polypeptides, agonists and antagonists are useful for
CC treating or diagnosing a cardiovascular, endothelial or angiogenic
CC disorder in a mammal, e.g. cardiac hypertrophy, trauma, cancer, age-
CC related macular degeneration, atherosclerosis, hypertension, arterial
CC restenosis, rheumatoid arthritis, angina, myocardial infarctions,
CC thrombophlebitis, lymphangitis, tumour angiogenesis (such as breast
CC carcinoma and liver carcinoma) and wound healing. The PRO nucleic acids
CC have applications in molecular biology, including use as hybridisation
CC probes, and in chromosome and gene mapping. This sequence encodes a novel
CC human secreted and transmembrane PRO polypeptide
XX
XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;
SQ Query Match 1.3%; Score 48; DB 7; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3715 AATGAAAAA
DB 1729 AATGAAAAA

RESULT 89

ABT44245
ID ABT44245 standard; cDNA; 1837 BP.

AC ABT44245;
XX
DT 06-NOV-2003 (first entry)
XX
DE Human PRO10111 cDNA.

XX PRO; proliferation; pericyte cell; TNF-alpha; blood; chondrocyte; ss;
KW differentiation; dermal fibroblast; tumour; gene therapy; gene;
KW cytotstatic.

XX Homo sapiens.
OS
XX
XX US2003050448-A1.

XX 13-MAR-2003.

XX 28-AUG-2002; 2002US-00230414.

XX 01-JUN-2001; 2001WO-US017800.
XX 29-JUN-2001; 2001WO-US021066.
XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX P-PSDB; ABJ72247.

XX New nucleic acid encoding for a PRO protein, useful for the manufacture
PT of a medicament for diagnosing or treating tumors or for measuring or
PT detecting expression of an associated gene.

XX Claim 2; Fig 39; 315pp; English.

XX The invention relates to a novel isolated nucleic acid encoding a fully
CC defined PRO polypeptide. The molecules of the invention may be useful for
CC stimulating proliferation or gene expression in pericyte cells or the
CC release of TNF-alpha from human blood. Other possible uses include the
CC stimulation or inhibition of chondrocyte proliferation or
CC differentiation, the stimulation of human dermal fibroblast cell
CC proliferation and the detection of the presence of a tumour within a
CC mammal. Furthermore, the nucleic acid may be useful for the manufacture
CC of a medicament for diagnosing or treating a tumour within a mammal or
CC for measuring or detecting the expression of an associated gene, as well
CC as during gene therapy. The current sequence is that of the human PRO
CC cDNA of the invention

XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
DB 1729 AATGAAAAA

RESULT 90

ABT44528
ID ABT44528 standard; cDNA; 1837 BP.

XX

AC ABT44528;

XX 06-NOV-2003 (first entry)

XX Human PRO10111 cDNA.

XX PRO; proliferation; gene; pericyte cell; TNF alpha; chondrocyte; blood;
KW tumour necrosis factor; proliferation; differentiation; gene therapy;
KW dermal fibroblast; ss.

XX Homo sapiens.

XX US2003027988-A1.

XX 06-FEB-2003.

XX 26-AUG-2002; 2002US-00227884.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
PI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX P-PSDB; ABJ72375.

XX New PRO protein encoding nucleic acid, useful for preparing PRO

PT polypeptides and anti-PRO antibodies for detecting the presence of a
PT tumor in a mammal.

XX Claim 2; Fig 39; 324pp; English.

XX The invention relates to a novel isolated PRO protein encoding nucleic
CC acid. The nucleic acid of the invention may be useful for preparing PRO
CC polypeptides and anti-PRO antibodies for detecting the presence of a
CC tumour in a mammal. Furthermore, the molecules of the invention may be
CC useful for stimulating proliferation or gene expression in pericyte
CC cells, the release of tumour necrosis factor (TNF)-alpha from human
CC blood, the proliferation or differentiation of chondrocyte cells and for
CC inhibiting the proliferation of normal human dermal fibroblast cells.
CC Finally, the molecules may be utilised during gene therapy. The current
CC sequence is that of the human PRO cDNA of the invention

XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA
DB 1729 AATGAAAAA

RESULT 91

ACD82195
ID ACD82195 standard; cDNA; 1837 BP.

XX ACD82195;

XX 19-SEP-2003 (first entry)

XX Human secreted/transmembrane polypeptide PRO 10111 cDNA.

XX Human; ss; chondrocyte stimulation; TNF-alpha stimulation; gene therapy;
KW human dermal fibroblast stimulation; tumour; tissue typing; gene;
KW affinity purification.

XX Homo sapiens.

```
XX PN US2003044934-A1.
XX PF
XX PP
XX PR 06-MAR-2003.
XX PT
XX PS 28-AUG-2002; 2002US-00230338.
XX SC 01-JUN-2001; 2001WO-US017800.
XX SD 29-JUN-2001; 2001WO-US021066.
XX SE 09-APR-2002; 2002US-00119480.
XX SF
XX SG (GETH ) GENENTECH INC.
XX SH Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX SI Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX SJ WPI: 2003-492274/46.
XX SK P-PSDB; ABO34270.
XX SL
XX SM New transmembrane polypeptides and nucleic acids encoding the
XX SN polypeptides, useful in gene therapy, in chromosome identification, as
XX SO chromosome markers, or in generating probes.
XX SP Claim 2; Fig 39; 315pp; English.
XX SQ The invention relates to an isolated nucleic acid encoding a PRO
XX SC polypeptide. Nucleic acids that encode PRO can be used to generate either
XX SD transgenic animals or knock-out animals useful in developing and
XX SE screening of therapeutically useful reagents. The nucleic acids may also
XX SF be used in gene therapy for replacing defective gene, in chromosome
XX SG identification, as chromosome markers, or in generating probes to isolate
XX SH full length PRO cDNA. The PRO polypeptides are useful for chondrocyte
XX SI stimulation, TNF-alpha stimulation, human dermal fibroblasts stimulation
XX SJ and for detecting the presence of tumour in a mammal. The PRO
XX SK polypeptides are useful as molecular markers for protein electrophoresis
XX SL and the isolated nucleic acids may be used for recombinantly expressing
XX SM those markers. The PRO polypeptides and nucleic acids may also be used in
XX SN tissue typing. Anti-PRO antibodies are useful in diagnostic assays for
XX SO PRO and in affinity purification of PRO from recombinant cell culture or
XX SI natural sources. The present sequence represents cDNA encoding a human
XX SJ secreted/transmembrane PRO polypeptide
XX SK Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;
XX SL Query Match 1.3%; Score 48; DB 8; Length 1837;
XX SM Best Local Similarity 100.0%; Pred. No. 3.6e-07;
XX SG Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX SH
XX SI QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
XX SJ |||||||
XX SK 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
XX SL
XX SM RESULT 92
XX SG ABT43901
XX SH ID ABT43901 standard; cDNA; 1837 BP.
XX SI AC ABT43901;
XX SJ
XX SK 16-OCT-2003 (first entry)
XX SL Human membrane bound receptor/protein PRO10111 cDNA sequence.
XX SM Human; PRO; membrane bound protein; membrane bound receptor;
XX SG cell proliferation; cell migration; cell differentiation;
XX SH mitogenic factor; survival factor; cytotoxic factor;
XX SI differentiation factor; neuroepithelial hormone; cell receptor;
XX SJ receptor-ligand interaction; cytostatic; chondrocyte; tumour; gene; ss.
XX SK Homo sapiens.
XX SL OS
XX SM US2003065147-A1.
XX SJ
XX SK
```

```
PD 03-APR-2003.
XX PF
XX PP
XX PR 29-AUG-2002; 2002US-00232224.
XX PT
XX PS 28-JUL-1999; 99US-0146222P.
XX SC 24-FEB-2000; 2000WO-US005004.
XX SD 02-MAR-2000; 2000WO-US005841.
XX SE 01-JUN-2001; 2001WO-US017800.
XX SF 29-JUN-2001; 2001WO-US021066.
XX SG 09-APR-2002; 2002US-00119480.
XX SH (GETH ) GENENTECH INC.
XX SI Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX SJ Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX SK WPI: 2003-522018/49.
XX SL P-PSDB; ABJ72077.
XX SM
XX SN One hundred and twenty two nucleic acids encoding PRO polypeptides,
XX SO useful for the manufacture of a medicament for diagnosing or treating
XX SI tumor.
XX SJ Claim 2; Fig 39; 315pp; English.
XX SK This invention relates to one hundred and twenty two novel nucleic acids
XX SL encoding human PRO membrane bound proteins or receptors. Extracellular
XX SM proteins play important roles in the formation, differentiation and
XX SG maintenance of multicellular organisms. The fate of many individual cells
XX SH (for example proliferation, migration or differentiation) is typically
XX SI governed by information received from other cells and the immediate
XX SJ environment. The information is often transmitted by secreted
XX SK polypeptides (for example mitogenic factors, survival factors, cytotoxic
XX SL factors, differentiation factors, neurotrophins and hormones) which are
XX SM received and interpreted by diverse cell receptors or membrane bound
XX SG proteins. These membrane bound proteins and receptors may be of use as
XX SH pharmaceutical and diagnostic agents, such as in the blocking of receptor
XX SI ligand interactions. The current invention provides the amino acid
XX SJ sequences of novel human membrane bound receptors and proteins, along
XX SK with the cDNA sequences encoding them. The novel proteins of the
XX SL invention may have cytostatic activities through the stimulation of
XX SM chondrocytes. The nucleic acids of the invention may be useful for the
XX SG manufacture of a medicament for diagnosing or treating a tumour in a
XX SH mammal. In addition, they may be useful for measuring or detecting the
XX SI expression of a tumour associated gene. The present sequence is the cDNA
XX SJ sequence encoding a human PRO protein of the invention
XX SK Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;
XX SL Query Match 1.3%; Score 48; DB 8; Length 1837;
XX SM Best Local Similarity 100.0%; Pred. No. 3.6e-07;
XX SG Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
XX SH
XX SI QY 3715 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
XX SJ |||||||
XX SK 1729 AATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776
XX SL
XX SM RESULT 93
XX SG ADB83529
XX SH ID ADB83529 standard; cDNA; 1837 BP.
XX SI AC ADB83529;
XX SJ
XX SK 04-DEC-2003 (first entry)
XX SL Novel human secreted and transmembrane protein PRO10111 cDNA.
XX SM human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
XX SG vulnery; antiarthritic; pericyte cell proliferation;
XX SH pericyte cell differentiation; chondrocyte cell proliferation;
XX SI chondrocyte cell differentiation; tumour necrosis factor alpha release;
XX SJ (TNF)-alpha release; dermal fibroblast cell proliferation;
XX SK
```


CC in a test sample of cells taken from the mammal, and a control sample of
 CC normal cells of the same cell type, where a higher level of expression of
 CC the PRO polypeptides in the test sample as compared to the control sample
 CC is indicative of the presence of tumour in the mammal. The tumour is lung
 CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. (I) is useful as molecular weight markers, for tissue
 CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 CC useful for chromosome and gene mapping or gene therapy. (II) is useful
 CC for generating transgenic animals or knock-out animals which are useful
 CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;
 Best Local Similarity 100.0%; Pred. No. 3.6e-07;
 Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
 DB 1729 AATGAA 1776

RESULT 95

ID ADB73176 standard; cDNA; 1837 BP.

AC ADB73176;

DT 04-DEC-2003 (first entry)

DE Novel human secreted and transmembrane protein PRO10111 cDNA.

XX human, secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritic; pericyte cell proliferation;
 KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX Homo sapiens.

XX US2003096968-A1.

XX 22-MAY-2003.

XX 29-AUG-2002; 2002US-00232223.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-765525/72.

XX P-PSDB; ADB73177.

XX New isolated PRO polypeptides useful as molecular weight markers in
 PT protein electrophoresis, useful for tissue typing, and for treating
 PT arthritis and tumors.

XX Claim 2; Fig 39; 308pp; English.

XX The invention describes an isolated PRO (secreted and transmembrane)
 CC polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are

CC useful for stimulating the proliferation of or gene expression in
 CC pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
 CC for stimulating the proliferation or differentiation of chondrocyte
 CC cells. PRO357, PRO725, PRO1155, PRO1306 or PRO1419 polypeptide
 CC are useful for stimulating the release of tumour necrosis factor (TNF)-
 CC alpha from human blood. PRO982, PRO357, PRO1306, PRO1419, PRO214,
 CC PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
 CC PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1309,
 CC PRO1025, PRO1181, PRO1126, PRO1186, PRO1192, PRO1244, PRO1412,
 CC PRO1286, PRO1330, PRO1347, PRO1305, PRO1279, PRO1340, PRO1338,
 CC PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1760, PRO1567,
 CC PRO1887, PRO1928, PRO341, PRO1801, PRO4333, PRO3543, PRO3432,
 CC PRO940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
 CC stimulating the proliferation of normal human dermal fibroblasts cells.
 CC PRO1811, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
 CC PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
 CC inhibiting the proliferation of normal human dermal fibroblast cells. PRO
 CC polypeptides such as PRO6004, PRO4981, PRO7174, PRO5778, PRO4332, etc.,
 CC are useful for detecting the presence of tumour in a mammal which
 CC involves comparing the level of expression of the above PRO polypeptides
 CC in a test sample of cells taken from the mammal, and a control sample of
 CC normal cells of the same cell type, where a higher level of expression of
 CC the PRO polypeptides in the test sample as compared to the control sample
 CC is indicative of the presence of tumour in the mammal. The tumour is lung
 CC tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
 CC liver tumour. (I) is useful as molecular weight markers, for tissue
 CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
 CC useful for chromosome and gene mapping or gene therapy. (II) is useful
 CC for generating transgenic animals or knock-out animals which are useful
 CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
 CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
 CC sport injuries). This sequence encodes a human secreted and transmembrane
 CC PRO polypeptide.

XX Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 8; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAA 3762
 DB 1729 AATGAA 1776

RESULT 96

ID ADB78258

XX ADB78258 standard; cDNA; 1837 BP.

XX ADB78258;

XX 04-DEC-2003 (first entry)

XX Novel human secreted and transmembrane protein PRO10111 cDNA.

XX Human, secreted and transmembrane protein; PRO; gene; ss; cytostatic;
 KW vulnary; antiarthritic; pericyte cell proliferation;
 KW pericyte cell differentiation; chondrocyte cell proliferation;
 KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
 KW (TNF)-alpha release; dermal fibroblast cell proliferation;
 KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
 KW colon tumour; breast tumour; prostate tumour; rectal tumour;
 KW liver tumour; tissue typing; chromosome mapping; gene mapping;
 KW gene therapy.

XX Homo sapiens.

XX US2003092889-A1.

XX 15-MAY-2003.

XX 13-AUG-2002; 2002US-00219478.

ID ADB78012 standard; cDNA; 1837 BP.
AC ADB78012;
XX
DT 04-DEC-2003 (first entry)
XX
DE Novel human secreted and transmembrane protein PRO10111 cDNA.
XX
KW Human; secreted and transmembrane protein; PRO; gene; ss; cytostatic;
KW vulnary; antiarthritic; pericyte cell proliferation;
KW pericyte cell differentiation; chondrocyte cell proliferation;
KW chondrocyte cell differentiation; tumour necrosis factor alpha release;
KW (TNF)-alpha release; dermal fibroblast cell proliferation;
KW dermal fibroblast cell differentiation inhibitor; tumour; lung tumour;
KW colon tumour; breast tumour; prostate tumour; rectal tumour;
KW liver tumour; tissue typing; chromosome mapping; gene mapping;
KW gene therapy.
XX
OS Homo sapiens.
XX
XX US2003092886-A1.
XX
XX 15-MAY-2003.
XX
XX 09-AUG-2002; 2002US-00216165.
XX
XX 25-JUL-2000; 2000US-0220607P.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-765494/72.
XX
XX P-PSDB; ADB78013.
XX
XX Novel isolated PRO polypeptide useful for tissue typing, gene therapy, as
XX molecular weight markers in protein electrophoresis, for treating
XX arthritis, tumor.
XX
XX Claim 2; Fig 39; 308pp; English.
XX
XX The invention describes an isolated PRO (secreted and transmembrane)
XX polypeptide (I). PRO982, PRO1160, PRO1187 or PRO1329 polypeptide are
XX useful for stimulating the proliferation of or gene expression in
XX pericyte cells. PRO357, PRO229, PRO1272 or PRO4405 polypeptide are useful
XX for stimulating the proliferation or differentiation of chondrocyte
XX cells. PRO231, PRO357, PRO725, PRO155, PRO1306 or PRO1419 polypeptide
XX are useful for stimulating the release of tumour necrosis factor (TNF)-
XX alpha from human blood. PRO982, PRO357, PRO725, PRO1306, PRO1419, PRO214,
XX PRO247, PRO337, PRO526, PRO363, PRO531, PRO1083, PRO840, PRO1080,
XX PRO1478, PRO1134, PRO826, PRO1005, PRO809, PRO1071, PRO1411, PRO1309,
XX PRO1025, PRO1124, PRO1126, PRO1186, PRO1192, PRO1274, PRO1412,
XX PRO1285, PRO1330, PRO1347, PRO1305, PRO1273, PRO1279, PRO1340, PRO1338,
XX PRO1343, PRO1376, PRO1387, PRO1409, PRO1474, PRO1917, PRO1567,
XX PRO1887, PRO1928, PRO1341, PRO1801, PRO4333, PRO3444, PRO4332,
XX PRO9940, PRO6079, PRO9836 or PRO10096 polypeptide are useful for
XX stimulating the proliferation of normal human dermal fibroblasts cells.
XX PRO181, PRO229, PRO788, PRO1194, PRO1272, PRO1488, PRO4302, PRO4408,
XX PRO5723, PRO5725, PRO7154, or PRO7425 polypeptide are useful for
XX inhibiting the proliferation of normal human dermal fibroblast cells. PRO
XX polypeptides such as PRO6004, PRO4981, PRO1714, PRO5778, PRO4332, etc.,
XX are useful for detecting the presence of tumour in a mammal which
XX involves comparing the level of expression of the above PRO polypeptides
XX in a test sample of cells taken from the mammal, and a control sample of
XX normal cells of the same cell type, where a higher level of expression of
XX the PRO polypeptides in the test sample as compared to the control sample
XX is indicative of the presence of tumour in the mammal. The tumour is lung
XX tumour, colon tumour, breast tumour, prostate tumour, rectal tumour or
XX liver tumour. (I) is useful as molecular weight markers, for tissue

CC typing, or as therapeutic agents. A polynucleotide (II) encoding (I) is
CC useful for chromosome and gene mapping or gene therapy. (II) is useful
CC for generating transgenic animals or knock-out animals which are useful
CC screening useful reagents. PRO357, PRO229, PRO1272 or PRO4405 polypeptide
CC is useful for treating bone and/or cartilage disorders (e.g., arthritis,
CC sport injuries). This sequence encodes a human secreted and transmembrane
CC PRO polypeptide.
XX
SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;
Query Match 1.3%; Score 48; DB 9; Length 1837;
Best Local Similarity 100.0%; Pred. No. 3.6e-07;
Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3715 AATGAAAAA
DB 1729 AATGAAAAA
RESULT 99
ADB78012
ID ADB78012 standard; cDNA; 1837 BP.
XX
XX ADB78012;
XX
XX 04-DEC-2003 (first entry)
XX
XX Human PRO polynucleotide #20.
XX
XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide;
KW tumour; cancer; lung; colon; breast; prostate; rectum; liver;
KW tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell;
KW pericyte cell; dermal fibroblast; bone disorder; cartilage disorder;
KW arthritis; sports injury; cytostatic; antiarthritic.
XX
XX Homo sapiens.
XX
XX US2003088067-A1.
XX
XX 08-MAY-2003.
XX
XX 13-AUG-2002; 2002US-00219479.
XX
XX 01-JUN-2001; 2001WO-US017800.
XX
XX 29-JUN-2001; 2001WO-US021066.
XX
XX 09-APR-2002; 2002US-00119480.
XX
XX (GETH) GENENTECH INC.
XX
XX Baker KP, Desnoyers L, Gerritsen ME, Goddard A, Godowski PJ;
XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;
XX WPI; 2003-657981/62.
XX
XX P-PSDB; ADB87079.
XX
XX One hundred and twenty two nucleic acids encoding PRO polypeptides,
XX useful in gene therapy, chromosome identification, tissue typing, or as
XX hybridization probes in chromosome and gene mapping.
XX
XX Claim 2; Fig 39; 314pp; English.
XX
XX The invention relates to human PRO polypeptides (secreted and
XX transmembrane polypeptides) and the PRO polynucleotides encoding them.
XX The PRO polypeptides and polynucleotides are useful as pharmaceuticals,
XX diagnostics, biosensors or bioreactors. They are particularly useful for
XX detecting tumours (e.g. lung tumour, colon tumour, breast tumour,
XX prostate tumour, rectal tumour or liver tumour) in a mammal, for
XX stimulating the release of tumour necrosis factor (TNF)-alpha from human
XX blood, for stimulating the proliferation or differentiation of
XX chondrocyte cells, for stimulating the proliferation of or gene
XX expression in pericyte cells or for stimulating the proliferation of
XX normal human dermal fibroblasts. The PRO nucleic acids are useful as
XX hybridisation probes, in chromosome and gene mapping, in generating

antisense RNA and DNA, in preparing PRO polypeptides by recombinant technology, in generating transgenic animals or knock-out animals which may be used in the development and screening of therapeutically useful reagents, in gene therapy, in chromosome identification, as chromosome markers, and in generating probes. The PRO polypeptides, or anti-PRO antibodies, are useful for preparing a medicament for treating a condition which is responsive to the PRO polypeptides or anti-PRO antibodies, such as pericyte-associated tumours and bone and/or cartilage disorders (e.g. arthritis, sports injuries), involving inducing the re-differentiation of chondrocytes. The PRO polypeptides are useful as molecular markers for protein electrophoresis, and in tissue typing. This sequence represents a human PRO polynucleotide of the invention.

XX SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 9; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

DB 1729 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

RESULT 100

ADB84660
ID ADB84660 standard; cDNA; 1837 BP.

AC ADB84660;

XX 04-DEC-2003 (first entry)

XX Human PRO polynucleotide #20.

XX Human; PRO; gene; ss; secreted polypeptide; transmembrane polypeptide; tumour; cancer; lung; colon; breast; prostate; rectum; liver; tumour necrosis factor-alpha; TNF-alpha; blood; chondrocyte cell; pericyte cell; dermal fibroblast; bone disorder; cartilage disorder; arthritis; sports injury; cytostatic; antiarthritic.

XX Homo sapiens.

XX US2003092890-A1.

XX 15-MAY-2003.

XX 14-AUG-2002; 2002US-00219536.

XX 28-JUL-1999; 99US-0146222P.

XX 24-FEB-2000; 2000WO-US005004.

XX 02-MAR-2000; 2000WO-US005841.

XX 01-JUN-2001; 2001WO-US017800.

XX 29-JUN-2001; 2001WO-US021066.

XX 09-APR-2002; 2002US-00119480.

XX (GETH) GENENTECH INC.

XX Baker KP, Deenoyers L, Gerritsen ME, Goddard A, Godowski PJ;

XX Grimaldi JC, Gurney AL, Smith V, Stephan JF, Watanabe CK, Wood WI;

XX WPI; 2003-777259/73.

XX P-PSDB; ADB84661.

XX New isolated PRO polypeptides, useful for tissue typing, gene therapy, as molecular weight markers in protein electrophoresis, and for treating arthritis and tumors.

XX Claim 2; Fig 39; 308pp; English.

XX The invention relates to human PRO polypeptides (secreted and transmembrane polypeptides) and the PRO polynucleotides encoding them. The PRO polypeptides and polynucleotides are useful as pharmaceuticals, diagnostics, biosensors or bioreactors. They are particularly useful for

CC detecting tumours (e.g. lung tumour, colon tumour, breast tumour, CC prostate tumour, rectal tumour or liver tumour) in a mammal, for CC stimulating the release of tumour necrosis factor (TNF)-alpha from human CC blood, for stimulating the proliferation or differentiation of CC chondrocyte cells, for stimulating the proliferation of or gene CC expression in pericyte cells or for stimulating the proliferation of CC normal human dermal fibroblasts. The PRO nucleic acids are useful as CC hybridisation probes, in chromosome and gene mapping, in generating CC antisense RNA and DNA, in preparing PRO polypeptides by recombinant CC technology, in generating transgenic animals or knock-out animals which CC may be used in the development and screening of therapeutically useful CC reagents, in gene therapy, in chromosome identification, as chromosome CC markers and in generating probes. The PRO polypeptides, or anti-PRO CC antibodies, are useful for preparing a medicament for treating a CC condition which is responsive to the PRO polypeptides or anti-PRO CC antibodies, such as pericyte-associated tumours and bone and/or cartilage CC disorders (e.g. arthritis, sports injuries), involving inducing the re- CC differentiation of chondrocytes. The PRO polypeptides are useful as CC molecular markers for protein electrophoresis, and in tissue typing. This CC sequence represents a human PRO polynucleotide of the invention.

XX SQ Sequence 1837 BP; 527 A; 454 C; 465 G; 391 T; 0 U; 0 Other;

Query Match 1.3%; Score 48; DB 9; Length 1837;

Best Local Similarity 100.0%; Pred. No. 3.6e-07;

Matches 48; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3715 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762

DB 1729 AATGAAAAA AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1776

Search completed: April 23, 2004, 04:26:38

Job time : 1354 secs.

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 23, 2004, 04:04:06 ; Search time 8944 Seconds
(without alignments)

12560.546 Million cell updates/sec

Title: -US-10-005-907-1

Perfect score: 3762

Sequence: 1 ggagaccgagtcactgtga.....aaaaaaaaaaaaaaaaaaaaa 3762

Scoring table: OLIGO_NUC
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size: 0

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 150 summaries

Database: EST*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_estci:*

9: gb_estci:*

10: gb_estci:*

11: gb_estci:*

12: gb_estci:*

13: gb_estci:*

14: gb_estci:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vit:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_phg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	665	17.7	741	BU155309	BU155309 AGENCOURT
2	624	16.6	785	BU171872	BU171872 AGENCOURT
3	620	16.5	782	BO441475	BO441475 AGENCOURT
4	602	16.0	728	CD639433	CD639433 AGENCOURT

ALIGNMENTS

[illegible]

Db 483 CTTTCCATATTTATTTTGTGTGTGAAGGCTCTTTTAGTCATATATAATTTTGGAT 542
 Qy 3569 TCATCTATGTTTAAATGTTCTATCAGTAGTGTGTACATCTTACTTGTCTCAGCATATCACCA 3628
 Db 543 TCATCTATGTTTAAATGTTCTATCAGTAGTGTGTACATCTTACTTGTCTCAGCATATCACCA 602
 Qy 3629 TATAGATATACTATAATTTG 3648
 Db 603 TATAGATATACTATAATTTG 622

 RESULT 4
 CD639433 728 bp mRNA linear EST 17-JUN-2003
 LOCUS AGENCOURT 14532881 NIH_MGC_191 Homo sapiens cDNA clone
 DEFINITION IMAGE:304114375 5', mRNA sequence.
 ACCESSION CD639433
 VERSION CD639433.1 GI:31806240
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 728) http://mgc.nci.nih.gov/
 TITILE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Daniela S. Gerhard, Ph.D.
 Office of Cancer Genomics
 National Cancer Institute
 Bldg. 31 Rm10A07 Bethesda, MD 20892
 Email: cgapbs-i@mail.nih.gov
 Tissue Procurement: Narayan Bhat
 CDNA Library Preparation: CLONTECH Laboratories, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: NDCM210 row: n column: 16
 High quality sequence stop: 604.
 Location/Qualifiers
 1. 728
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:304114375"
 /tissue_type="Pooled"
 /lab_host="DH10B (T1 phage-resistant)"
 /clone_lib="NIH_MGC_191"
 /note="vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc);
 Site2: SfiI (ggcgctctggcc); Library is oligo-dT primed
 and directionally cloned. PBMC - Peripheral Blood
 Mononuclear Cells. RNA was pooled from 3/6hour stimulation
 with PMA adn Ionomycin. 5' and 3' adaptors were used in
 cloning as follows: 5' adaptor sequence:
 5'-CAGCGCCATTATGCCC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCCGAGGGCGGCACATG-T(30)BN-3' (where B = A,
 C, G and N = A, C, G, or T). Average insert size 1.69
 kb (range 0.70-5.0 kb). 15/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

 ORIGIN
 Query Match 16.0%; Score 602; DB 14; Length 728;
 Best Local Similarity 100.0%; Pred. No. 4e-105;
 Matches 602; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

 Qy 1835 ATCTTACCGAAGTGATGTGTAGTTCTCGTTTAAATTCAGCAAACTGAAAATA 1894
 Db 3 ATCTTACCGAAGTGATGTGTAGTTCTCGTTTAAATTCAGCAAACTGAAAATA 62

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 704)
AUTHORS Zhao,S., Adams,M.D., Nierman,W., Malek,J., Shizuya,H., Simon,M. and Venter,J.C.
TITLE Use of BAC End Sequences from Caltech Libraries for Sequence-Ready Map Building
JOURNAL Unpublished (1997)
COMMENT Other_GSSs: CITBI-EL-2548B21.TR
Contact: Shaying Zhao, William Nierman, Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850
Tel: 301 838 0200
Fax: 301 838 0208
Email: hbe@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC end search page:
http://www.tigr.org/tldb/hungen/bac_end_search/bac_end_search.html.
Seq primer: M3-21
Class: BAC ends.

FEATURES
source
1..704
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="2548B21"
/sex="male"
/cell_type="sperm"
/clone_lib="CITBI-EL"
/note="Vector: pBeloBAC11; Site_1: EcoRI; Site_2: EcoRI; Caltech Human BAC Library D"

ORIGIN
Query Match 14.8%; Score 556; DB 28; Length 704;
Best Local Similarity 99.7%; Pred. No. 2e-96;
Matches 656; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 803 GAAATTCGGCCATGATAGAGAGGAGGTGACACACACTTGTATACCCCTTCCCTTTG 862
DB 6 GAAATTCGGCCATGATAGAGAGGAGGTGACACACACTTGTATACCCCTTCCCTTTG 65
QY 863 GAGTTATGCACAGTGACAGGATGAGTCTATAGACTGATGAATAGACTGATGGC 922
DB 66 GAGTTATGCACAGTGACAGGATGAGTCTATAGACTGATGAATAGACTGATGGC 125
QY 923 AATAGAGTCCCAATTCACACTGCTGCTGTAGATCACACACTGTCTGAGGATTC 982
DB 126 AATAGAGTCCCAATTCACACTGCTGCTGTAGATCACACACTGTCTGAGGATTC 185
QY 983 ATCTATGAGCTTTGTCTACATAACAGAGACCTTGGTTTCCAAACCCCTTATTTAGC 1042
DB 186 ATCTATGAGCTTTGTCTACATAACAGAGACCTTGGTTTCCAAACCCCTTATTTAGC 245
QY 1043 TAAAGCATTTTCTACTGACTCTTAAGCTTTAGCAAGCTTAACCTTTCAACCA 1102
DB 246 TAAAGCATTTTCTACTGACTCTTAAGCTTTAGCAAGCTTAACCTTTCAACCA 305
QY 1103 ATTGCCAATCAGACAACTTTGAATCTACCTATGACCTGTGAAGCTCTCTCTGCTCAAG 1162
DB 306 ATTGCCAATCAGACAACTTTGAATCTACCTATGACCTGTGAAGCTCTCTCTGCTCAAG 365
QY 1163 ATCTTGCTCTTAAAGCTGAACCATGCACTTCCATTTAATGATTTATGCTTTGCT 1222
DB 366 ATCTTGCTCTTAAAGCTGAACCATGCACTTCCATTTAATGATTTATGCTTTGCT 425
QY 1223 TGTAACTCTGCTCCCTAAATGTATAAAGTAAACCGTGACCTGACCCACTCAGGCAC 1282
DB 426 TGTAACTCTGCTCCCTAAATGTATAAAGTAAACCGTGACCTGACCCACTCAGGCAC 485
QY 1283 ACTTTCTCAGGACCTCTGAGGTGATCCAGGCCATGGTAAAGTCATGTTGGCTCAGAA 1342

DB 486 ACTTCTCAGGACCTCCTGAGAGTGTATCCAGGCGCATGGTAAGTCATGTTGGCTCAGAA 545
QY 1343 TCAACCTCTTTAAATATTTTACAGAAATTTGGTTTGGTTTACCAATAAGTCTCCCAAAAT 1402
DB 546 TCAACCTCTTTAAATATTTTACAGAAATTTGGTTTGGTTTACCAATAAGTCTCCCAAAAT 605
QY 1403 ATATGTCCAGAAATCTTCAATTCAGGCTGCTCACCAGAAATTTCAATGCCAATCT 1460
DB 606 ATATGTCCAGAAATCTTCAATTCAGGCTGCTCACCAGAAATTTCAATGCCAATCT 663

RESULT 7
BF434786/c
LOCUS BF434786.1
DEFINITION mRNA sequence.
ACCESSION BF434786
VERSION BF434786.1
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 549)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R. Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Sequencing by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL, send email to: info@image.llnl.gov
Seq primer: -400P from Gibco
High quality sequence stop: 476.

FEATURES
source
1..549
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:3641981"
/lab_host="DH10B"
/clone_lib="NCI CGAP Kid11"
/note="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; Plasmid DNA from the normalized library NCI_CGAP Kid3 was prepared, and ss circles were made in vitro. Following HAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from a pool of 5,000 clones made from the same library (cloneID# 132376-132391, 1456007-1456775, and 1500552-1502855). Subtraction by Bento Soares and M. Fatima Bonaldo."

ORIGIN
Query Match 14.1%; Score 530; DB 10; Length 549;
Best Local Similarity 100.0%; Pred. No. 2.1e-91;
Matches 530; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3191 TCGTTTTTAACTGTAAATTTTGTGAAGCTTATCTTTATGCATATAATATTTCACAT 3250
DB 549 TCGTTTTTAACTGTAAATTTTGTGAAGCTTATCTTTATGCATATAATATTTCACAT 490
QY 3251 TTATACATCTTTTATATTTTAAATCAGTTTTTACTCAAGTGTGATTATATACAGAAATGT 3310
DB 489 TTATACATCTTTTATATTTTAAATCAGTTTTTACTCAAGTGTGATTATATACAGAAATGT 430
QY 3311 AACCACTGTAAGGGTAGAGTTATAGAAATTTTGTCAAATGTATTCCACCATGTAGTCACC 3370

```

Db 429 AACCACTGTAGGTTATAGAAATTTTGCAATGATATCCACCAATGATGACACC 370
QY 3371 TCCTTATGAAGAGACAGAACGCTACATCTCCAGAAAGTTCACAGTCTCTTTTCC 3430
Db 369 TCCTTATGAAGAGACAGAACGCTACATCTCCAGAAAGTTCACAGTCTCTTTTCC 310
QY 3431 CTGAGTTTCCACGCTCTCGCAACCAATGATCTCTTCGTATATATATACTGTTCTAGA 3490
Db 309 CTGAGTTTCCACGCTCTCGCAACCAATGATCTCTTCGTATATATATACTGTTCTAGA 250
QY 3491 TATTTGAGCAATGACCTTTCCATATTTATTTTGTGTGTAAGCTTCTTTTGTCTA 3550
Db 249 TATTTGAGCAATGACCTTTCCATATTTATTTTGTGTGTAAGCTTCTTTTGTCTA 190
QY 3551 TTATAATATTTTTCAGATTCATCTGTTTAAATCTCTATCAGTAGTGTGACATCTTACT 3610
Db 189 TTATAATATTTTTCAGATTCATCTGTTTAAATCTCTATCAGTAGTGTGACATCTTACT 130
QY 3611 TGTCTCAGCATATCACCATATAGATATATATAATTTTGTAACTAATCTATCTGATGATA 3670
Db 129 TGTCTCAGCATATCACCATATAGATATATATAATTTTGTAACTAATCTATCTGATGATA 70
QY 3671 TGTAGGATATTTAGTTTTCACATTTATGAATTAAGTGGCTATATAATGAA 3720
Db 69 TGTAGGATATTTAGTTTTCACATTTATGAATTAAGTGGCTATATAATGAA 20

RESULT 8
BF242113 527 bp mRNA linear EST 14-NOV-2000
LOCUS 601880401F1 NIH_MGC_55 Homo sapiens CDNA clone IMAGE:4109064 5',
DEFINITION mRNA sequence.
ACCESSION BF242113
VERSION BF242113.1 GI:11156040
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 527)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: c9apbs-r@mail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: CLONETECH Laboratories, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW993 row: a column: 01
High quality sequence stop: 518.
Location/Qualifiers
1. 527
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4109064"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 55"
/notes="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site 1: SfiI (ggcgctcgcc); Site 2: SfiI
(ggcatatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and
3' adaptor sequence:
5'-ATTCTAGAGCGCGAGCGCGGCACATG-dT(30)BN-3' (where B = A,
C, or G and N = A, C, G, or T). Average insert size
1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained

```

inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

```

Query Match 13.5%; Score 506; DB 10; Length 527;
Best Local Similarity 100.0%; Pred. No. 7.4e-87;
Matches 506; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GAGAACCGAGTCACTGTGAAAGATGGAAATATATCTCTCGGAAACTCAGTTGCCGTG 60
Db 21 GAGAACCGAGTCACTGTGAAAGATGGAAATATATCTCTCGGAAACTCAGTTGCCGTG 80
QY 61 GGAGAGAAATCAAAAGAACCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAA 120
Db 81 GGAGAGAAATCAAAAGAACCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAA 140
QY 121 ATGACTACATTTTGAAGAAACTTCAAGATCAAGATTAAGAAAGCCAAAGTTTCATCC 180
Db 141 ATGACTACATTTTGAAGAAACTTCAAGATCAAGATTAAGAAAGCCAAAGTTTCATCC 200
QY 181 ACTTCTAATCAGAAACCCAGAAATGGCAGTGGTCTTGAAGAGTGTGCTACATCTCAT 240
Db 201 ACTTCTAATCAGAAACCCAGAAATGGCAGTGGTCTTGAAGAGTGTGCTACATCTCAT 260
QY 241 AATCAGATCCCCCATCAGAGATCTCCCTGAGCTCAATGATGATGGCTATGAGACATT 300
Db 261 AATCAGATCCCCCATCAGAGATCTCCCTGAGCTCAATGATGATGGCTATGAGACATT 320
QY 301 GACTCCCTCACAAGGAAAGTGAAGAGTGTAGAGAAAGTTCAGACACAGAAATATGCCCTT 360
Db 321 GACTCCCTCACAAGGAAAGTGAAGAGTGTAGAGAAAGTTCAGACACAGAAATATGCCCTT 380
QY 361 CTTAGAGCTTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATTAAGTTGTG 420
Db 381 CTTAGAGCTTCTGTAGTAGGCTTGTCTCTGACCCATGAGCATGATTAAGTTGTG 440
QY 421 TTTCACACATAAATCTCAAGCTGCTTTATCACCTTCCAGCAATGAACAAATGCAGAA 480
Db 441 TTTCACACATAAATCTCAAGCTGCTTTATCACCTTCCAGCAATGAACAAATGCAGAA 500
QY 481 TAGCAGACTCTGGCGAAGTTGTTCAC 506
Db 501 TAGCAGACTCTGGCGAAGTTGTTCAC 526

```

RESULT 9

```

BF2600841 757 bp DNA linear GSS 08-JUN-2003
LOCUS WHAAJ89TR Human MCF7 breast cancer cell line library (MCF7_1) Homo
DEFINITION sapiens genomic clone MCF7_1-3010, genomic survey sequence.
ACCESSION BF2600841
VERSION BF2600841.1 GI:31509303
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 757)
AUTHORS Volik S., Zhao S., Chin K., Brebner, J.H., Herndon, D.R., Tao, Q.,
Kowbel, D., Huang G., Lapuk, A., Kuo, W.-L., Magrane, G., de Jong, P.,
Gray, J.W., and Collins, C.
TITLE End-sequence profiling: Sequence-based analysis of aberrant genomes
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 100 (13), 7696-7701 (2003)
MEDLINE 22709111
PUBMED 12788976
COMMENT Contact: Volik SV
Colin Collins' lab
UCSF Comprehensive Cancer Center
UCSF Box 0808, San Francisco, CA 94143-0808, USA
Tel: 415 502 7066
Fax: 415 502 5665
Email: svolik@cc.ucsf.edu

```

FEATURES source

This clone is available from Amplicon Express
http://www.genomex.com
Class: BAC ends.

FEATURES

source
Location/Qualifiers
1..757
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone="MCF7_1-3010"
/sex="female"
/clone_lib="Human MCF7 breast cancer cell line library (MCF7_1)"
/note="Vector: pECBAC1; Site 1: HindIII; This library was constructed from MCF7 breast cancer cell line by Amplicon Express (http://www.genomex.com) using their standard procedure."

ORIGIN

Query Match 12.5%; Score 471; DB 28; Length 757;
Best Local Similarity 100.0%; Pred. No. 2.1e-80;
Matches 471; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3222 ATCTTTTATGCATATAATATTTGAACATTTTACATTTGTTTATATTTTATTCAGTTTGA 3281
Db 1 ATCTTTTATGCATATAATATTTGAACATTTTACATTTGTTTATATTTTATTCAGTTTGA 60
QY 3282 CTCAGTGTGATTATATACAGAAATGTAACCACTGTAGGCTAGAGTTATAGAAATTT 3341
Db 61 CTCAGTGTGATTATATACAGAAATGTAACCACTGTAGGCTAGAGTTATAGAAATTT 120
QY 3342 TGTCAATGTATTCACCATGTAGTCACCTCCTTATGAAGAGACAGACAGTACATCCT 3401
Db 121 TGTCAATGTATTCACCATGTAGTCACCTCCTTATGAAGAGACAGACAGTACATCCT 180
QY 3402 CCCAGAAAGTCCACAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGGAAACCAATGAT 3461
Db 181 CCCAGAAAGTCCACAGTGTCTCTTTCCCTGAGTTTCCACAGTCTCTGGGAAACCAATGAT 240
QY 3462 CTGCTTCGTATATAATTAATCTCTAGATATTTGTAGCAATGTACCTTTCCATATTTA 3521
Db 241 CTGCTTCGTATATAATTAATCTCTAGATATTTGTAGCAATGTACCTTTCCATATTTA 300
QY 3522 TTTTGTGTGTGAAGCTCTTTTGTAGTCATTAATATTTTGTAGATTCATCTATGTTTA 3581
Db 301 TTTTGTGTGTGAAGCTCTTTTGTAGTCATTAATATTTTGTAGATTCATCTATGTTTA 360
QY 3582 ATGTTCTATCATGATGTGTACATCTTCTGTCTCAGCATATCACCATATAGATATACTA 3641
Db 361 ATGTTCTATCATGATGTGTACATCTTCTGTCTCAGCATATCACCATATAGATATACTA 420
QY 3642 TAATTTGTTTAACTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 3692
Db 421 TAATTTGTTTAACTCACTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 471

RESULT 10

BU155429
LOCUS BU155429 786 bp mRNA linear EST 03-SEP-2002
DEFINITION AGENCOURT 7979368 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:6108338
5', mRNA sequence.
ACCESSION BU155429
VERSION BU155429.1 GI:22668961
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 786)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
CONTACT: Robert Strausberg, Ph.D.
Email: cgaabs-romail.nih.gov

Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at:
http://image.llnl.gov
Plate: LLCM2352 row: h column: 03
High quality sequence stop: 399.

FEATURES

source
Location/Qualifiers
1..786
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6108338"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGCGCGCGCCGACATG-DT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 12.3%; Score 462; DB 13; Length 786;
Best Local Similarity 99.8%; Pred. No. 1e-78;
Matches 512; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2609 GTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTAGGATTTTCAACTGTACTGACCT 2668
Db 3 GTTTCAGATTTTGGAGCATTTTCAGATCTTCAGATTAGGATTTTCAACTGTACTGACCT 62
QY 2669 TTTAGTCATTGACAGCATTAATCAATAGTGGATCCAGATACTCATTTTGTCTGTATAC 2728
Db 63 TTTAGTCATTGACAGCATTAATCAATAGTGGATCCAGATACTCATTTTGTCTGTATAC 122
QY 2729 ACATTTTGGCTCTCTATTCAACGAATTTCTATGCCCTCTTGTGTGATTTTAAATGTGGG 2788
Db 123 ACATTTTGGCTCTCTATTCAACGAATTTCTATGCCCTCTTGTGTGATTTTAAATGTGGG 182
QY 2789 AAGGGAACAATAGAAATTTTGCATTTCTAGAAAGTCATTTCTGTCAAAATATGTCATC 2848
Db 183 AAGGGAACAATAGAAATTTTGCATTTCTAGAAAGTCATTTCTGTCAAAATATGTCATC 242
QY 2849 CTGTAGATATTAGCCAAATTTTAGGAAATGACAAAATTTTCTGCTGTGCTTTG 2908
Db 243 CTGTAGATATTAGCCAAATTTTAGGAAATGACAAAATTTTCTGCTGTGCTTTG 302
QY 2909 TAGCTCTTTTATGATATAATTAATACCTTTTGTATAAATTAATTTTAAATTTAGTACAC 2968
Db 303 TAGCTCTTTTATGATATAATTAATACCTTTTGTATAAATTAATTTTAAATTTAGTACAC 362
QY 2969 ATCTGGAATTTATCAGAGAAGGGGCAAGCAATAGGTTTAAATAGATTTTGTAGTAA 3028
Db 363 ATCTGGAATTTATCAGAGAAGGGGCAAGCAATAGGTTTAAATAGATTTTGTAGTAA 422
QY 3029 GGAACCTTGAATTCAGAGCATCAATGTCTTCTGTGTGTTTCCATTAAGCCACAGCAGA 3088
Db 423 GGAACCTTGAATTCAGAGCATCAATGTCTTCTGTGTGTTTCCATTAAGCCACAGCAGA 482
QY 3089 TGCTTTAACTTCTCCGAGATCTAGTTTTTTCAGG 3121
Db 483 TGCTTTAACTTCTCCGAGATCTAGTTTTTTCAGG 515

```
RESULT 11
AW65849/c 504 bp mRNA linear EST 06-APR-2000
LOCUS h194f08.x1 Soares_NFL_T_GBC_S1 Homo sapiens cDNA clone
DEFINITION IMAGE:2979975 3', mRNA sequence.
ACCESSION AW65849
VERSION AW65849.1 GI:7458398
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 504)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40UP from Gibco
High quality sequence stop: 451.
FEATURES
source
1..504
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2979975"
/lab_host="DH10B"
/clone_lib="Soares_NFL_T_GBC_S1"
/notes="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Equal amounts of plasmid DNA from three normalized
libraries (fetal lung NBHL19, testis NHT, and B-cell
NCI CGAP_GCB) were mixed and ss circles were made in
vitro. Following HAP purification, this DNA was used as
tracer in a subtractive hybridization reaction. The driver
was PCR-amplified cDNAs from pools of 5,000 clones made
from the same 3 libraries. The pools consisted of
I.M.A.G.E. clones 297480-302087, 682632-687239,
726408-728711, and 729096-731399. Subtraction by Bento
Soares and M. Fatima Bonaldo."
```

ORIGIN

```
Query Match 11.7%; Score 441; DB 10; Length 504;
Best Local Similarity 99.8%; Pred. No. 1.5e-74;
Matches 491; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 3219 CTTATCTTTTATGCATATAAATTTTGAACATTTTACATGTTTATATTTTATATCGATT 3278
Db 504 CTTATCTTTTATGCATATAAATTTTGAACATTTTACATGTTTATATTTTATATCGATT 445

Qy 3279 TTACTCAAGTGTGATTATATACAGAAATGTAACCACTGTAAGGGTAGAGTTATAGAA 3338
Db 444 TTACTCAAGTGTGATTATATACAGAAATGTAACCACTGTAAGGGTAGAGTTATAGAA 395

Qy 3339 TTTTGTCAAAATGTAATCCCATGTAGTCACCTCTTATGAGAGACAGAACACGTACAT 3398
Db 384 TTTTGTCAAAATGTAATCCCATGTAGTCACCTCTTATGAGAGACAGAACACGTACAT 325

Qy 3399 CTTCCAGAAAGTTCACAGTGCTCTTTTCCCTGAGTTTCACAGTCTCTGGCAACCAAT 3458
Db 324 CTTCCAGAAAGTTCACAGTGCTCTTTTCCCTGAGTTTCACAGTCTCTGGCAACCAAT 265

Qy 3459 GATCTGCTCGTATAAATTAATTAACCTGTTAGATATTTGTAGCAATAGACCTTTCCATAT 3518
Db 264 GATCTGCTCGTATAAATTAATTAACCTGTTAGATATTTGTAGCAATAGACCTTTACATAT 205

Qy 3519 TTATTTTGTGTGTAAGCGTCTCTTTTAGTCATTATATATTTTGTAGATTCATCTATGT 3578
Db 204 TTATTTTGTGTGTAAGCGTCTCTTTTAGTCATTATATATTTTGTAGATTCATCTATGT 145
```

```
Qy 3579 TTAATGTTCTATCAGTAGTTCATCATCTTCTCTCAGCATATCACCATATAGATATA 3638
Db 144 TTAATGTTCTATCAGTAGTTCATCATCTTCTCTCAGCATATCACCATATAGATATA 85

Qy 3639 CTATAAATTTGTTAATCTAATCACTGATGATATAGGATATTTAAGTTTTTGACATTAT 3698
Db 84 CTATAAATTTGTTAATCTAATCACTGATGATATAGGATATTTAAGTTTTTGACATTAT 25

Qy 3699 GAATTAAGTGCC 3710
Db 24 GAATTAAGTGCC 13

RESULT 12
AW66257/c 456 bp mRNA linear EST 20-DEC-1999
LOCUS wh86904.x1 NCI_CGAP_Kid1l Homo sapiens cDNA clone IMAGE:2385942 3',
DEFINITION mRNA sequence.
ACCESSION AW66257
VERSION AW66257.1 GI:5232766
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 456)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: M. Bento Soares, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert Length: 1086 Std Error: 0.00
Seq primer: -40UP from Gibco
High quality sequence stop: 437.
FEATURES
source
1..456
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2385942"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Kid1l"
/notes="Organ: Kidney; Vector: pT7T3D-Pac (Pharmacia) with
a modified polylinker; Site 1: Not I; Site 2: Eco RI;
Plasmid DNA from the normalized library NCI_CGAP_Kid3 was
prepared, and ss circles were made in vitro. Following HAP
purification, this DNA was used as tracer in a subtractive
hybridization reaction. The driver was PCR-amplified cDNAs
from a pool of 5,000 clones made from the same library
(cloneIDB 1322376-1323911, 1456007-1456775, and
1500552-1502855). Subtraction by Bento Soares and M.
Fatima Bonaldo."
```

ORIGIN

```
Query Match 11.6%; Score 438; DB 9; Length 456;
Best Local Similarity 100.0%; Pred. No. 6e-74;
Matches 438; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3283 TCAAGTGTGATTATATACAGAAATGTAACCACTGTAAGGGTAGAGTTATAGATTTT 3342
Db 456 TCAAGTGTGATTATATACAGAAATGTAACCACTGTAAGGGTAGAGTTATAGATTTT 397

Qy 3343 GTCAATGTATTCAACCATGTAGTCACCTCTTATGAGAGACAGAACACGTACATCTTC 3402
```

Db 396 GTCRAATGATATCCACCATGTAGTCACTCCTTATGAGAGACAGAACAGTACATCCTC 337
 QY 3403 CCAGAAAGTCCACAGTCTCTCTTTCCCTCAGTTTACCAAGTCTCTGCGACCAATGATC 3462
 Db 336 CCAGAAAGTCCACAGTCTCTCTTTCCCTCAGTTTACCAAGTCTCTGCGACCAATGATC 277
 QY 3463 TGCCTTCGTAATATTAATACTGTTCTAGATATTTCTAGCAATGTACCTTTCCCATATTTAT 3522
 Db 276 TGCCTTCGTAATATTAATACTGTTCTAGATATTTCTAGCAATGTACCTTTCCCATATTTAT 217
 QY 3523 TTTGTGTGTGAAGCTCTCTTTTGTAGTCAATTAATATTTTGTAGATCAATCTATGTTAA 3582
 Db 216 TTTGTGTGTGAAGCTCTCTTTTGTAGTCAATTAATATTTTGTAGATCAATCTATGTTAA 157
 QY 3583 TGTTCCTATCAGTAGTTGACATCTTACTTGTCTCAGCATATCACCATATAGATATCTAT 3642
 Db 156 TGTTCCTATCAGTAGTTGACATCTTACTTGTCTCAGCATATCACCATATAGATATCTAT 97
 QY 3643 AATTTGTTAACTAATCACTGATGATATGATGATATTTAAGTTTGTGACATATGAAAT 3702
 Db 96 AATTTGTTAACTAATCACTGATGATATGATGATATTTAAGTTTGTGACATATGAAAT 37
 QY 3703 AAAGTGGCTATAAATGAA 3720
 Db 36 AAAGTGGCTATAAATGAA 19

RESULT 13
 LOCUS BG461807/c 549 bp mRNA linear EST 21-APR-2001
 DEFINITION R5744735 Athersys RAGE Library Homo sapiens cDNA, mRNA sequence.
 ACCESSION BG461807
 VERSION BG461807.1 GI:13750313
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 549)
 AUTHORS Harrington,J.J., Sherf,B., Rundlett,S., Jackson,P.D., Perry,R.,
 Cain,S., Leventhal,C., Thornton,M., Ramachandran,K.,
 Whittington,J., Lerner,L., Costanzo,D., McElligott,K., Booser,S.,
 Mays,R., Smith,E., Veloso,N., Klika,A., Hess,J., Cothren,K., Lo,K.,
 Offenbacher,J., Danzig,J. and Ducar,M.
 * TITLE Creation of genome-wide protein expression libraries using random
 activation of gene expression
 Nat. Biotechnol. 19 (5), 440-445 (2001)
 JOURNAL 21227151
 MEDLINE 11329013
 PUBMED
 COMMENT Contact: Scott J. Cain
 Athersys, Inc.
 3201 Carnegie Ave, Cleveland, OH 44115, USA
 Tel: 216 431 9900
 Fax: 216 361 9596
 Email: scain@athersys.com.

FEATURES
 source
 1..549
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /cell_line="HT1080"
 /clone_lib="Athersys RAGE Library"
 /note="See 'Creation of Genome-wide Protein Expression
 Libraries using Random Activation of Gene Expression',
 Nature Biotechnology, in press. Note that even though the
 cell type indicated is HT1080, since a random activation
 method was used, these sequence tags are not necessarily
 expressed in HT1080 under normal circumstances."

ORIGIN
 Query Match 10.3%; Score 387; DB 12; Length 549;
 Best Local Similarity 100.0%; Pred. NO. 2.1e-64;
 Matches 387; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3316 CTGTAAGGTAGAGTATAAGAAATTTTGTCAAATGATTTCAACCATGTAGTACACCTCCTT 3375
 Db 387 CTGTAAGGTAGAGTATAAGAAATTTTGTCAAATGATTTTCAACCATGTAGTACACCTCCTT 328
 QY 3376 ATGAAGAGACAGAACACGCTACATCTCCAGAAAGTTCCACAGTCTCTCTTTTCCCTGAG 3435
 Db 327 ATGAAGAGACAGAACACGCTACATCTCCAGAAAGTTCCACAGTCTCTCTTTTCCCTGAG 268
 QY 3436 TTTTCAACGCTCTGCGCAACCATGATCTGCTTCGTATTAATTAATTAATTAATTAATTAATTT 3495
 Db 267 TTTTCAACGCTCTGCGCAACCATGATCTGCTTCGTATTAATTAATTAATTAATTAATTT 208
 QY 3496 GTAGCAATGTACCTTTCCATATTTTATTTTGTGTGTGAAGCTTCTTTTGTAGTCAATTATA 3555
 Db 207 GTAGCAATGTACCTTTCCATATTTTATTTTGTGTGTGAAGCTTCTTTTGTAGTCAATTATA 148
 QY 3556 ATATTTTGTAGATTCATCTATGTTTAAATGTTCTATCAGTAGTTGACATCTTACTTGTCT 3615
 Db 147 ATATTTTGTAGATTCATCTATGTTTAAATGTTCTATCAGTAGTTGACATCTTACTTGTCT 88
 QY 3616 CAGCATATCACCATATAGATATCTATTAATTTGTTAACTTAATCACTGATGATATGTAG 3675
 Db 87 CAGCATATCACCATATAGATATCTATTAATTTGTTAACTTAATCACTGATGATATGTAG 28
 QY 3676 GATATTTAAAGTTTGTGACATTAATGAAT 3702
 Db 27 GATATTTAAAGTTTGTGACATTAATGAAT 1

RESULT 14
 LOCUS AA447527/c 416 bp mRNA linear EST 04-JUN-1997
 DEFINITION zw81b09.sl Soares testis NHT Homo sapiens cDNA clone IMAGE:782585
 3', mRNA sequence.
 ACCESSION AA447527
 VERSION AA447527.1 GI:21611197
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 416)
 AUTHORS Hillier,L., Allen,M., Bowles,L., Dubuque,T., Geisel,G., Jost,S.,
 Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin,J.,
 Moore,B., Schellenberg,K., Steptoe,M., Tan,P., Theising,B.,
 White,Y., Wyllie,T., Waterston,R. and Wilson,R.
 WashU-Merck EST Project 1997
 Unpublished (1997)
 COMMENT Contact: Wilson RK
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@wustl.wustl.edu
 This clone is available royalty-free through LNL ; contact the
 IMAGE Consortium (info@image.llnl.gov) for further information.
 Seq primer: -41m13 fwd. Et from Amersham
 High quality sequence stop: 102.

FEATURES
 source
 1..416
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:782585"
 /sex="male"
 /lab_host="DH10B"
 /clone_lib="Soares testis NHT"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA
 was prepared from mRNA obtained from Clontech
 Laboratories, Inc., and primed with a Not I - oligo(dT)
 primer [5]

TGTTACCAATCTGAAGTGGAGCGCGCCCAATTTTTTTTTTTTTTTT 3'.
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified p773 vector. Library
went through one round of normalization to Cot5, and was
constructed by Bento Soares and M. Fatima Bonaldo. "

ORIGIN
Query Match 8.4%; Score 316; DB 9; Length 416;
Best Local Similarity 100.0%; Pred. No. 7.1e-51;
Matches 316; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 3405 AGAAGTTCACAGTGCCTTTTCCCTGAGTTTCACAGTCCCTGGCAACCAATGATCTG 3464
Db 343 AGAAGTTCACAGTGCCTTTTCCCTGAGTTTCACAGTCCCTGGCAACCAATGATCTG 284
Qy 3465 CTTGCTATAATTAATCACTGTTCTAGATATTGTAGCAATGATACCTTTCCATATTTTATTT 3524
Db 283 CTTGCTATAATTAATCACTGTTCTAGATATTGTAGCAATGATACCTTTCCATATTTTATTT 224
Qy 3525 TGTGCTGAAGGCTTCTTTTAGTCATTAATATTTTGTAGATTCATCTATGTTTAAAG 3584
Db 223 TGTGCTGAAGGCTTCTTTTAGTCATTAATATTTTGTAGATTCATCTATGTTTAAAG 164
Qy 3585 TTCTATCAGTAGTGTACATCTTACTTCTCAGCATATCACCATAGATATACTATAA 3644
Db 163 TTCTATCAGTAGTGTACATCTTACTTCTCAGCATATCACCATAGATATACTATAA 104
Qy 3645 TTTGTTAATCTAATCACTGATGATGATAGATATTTAAGTTTTCACATTAATGAATAA 3704
Db 103 TTTGTTAATCTAATCACTGATGATGATAGATATTTAAGTTTTCACATTAATGAATAA 44
Qy 3705 AGTGCTATAATGAATGA 3720
Db 43 AGTGCTATAATGAATGA 28

RESULT 15
AQ006627/c
LOCUS
DEFINITION
CIT-HSP-2294F11.TF CIT-HSP Homo sapiens genomic clone 2294F11,
genomic survey sequence.
ACCESSION
AQ006627
VERSION
AQ006627.1 GI:3081444
KEYWORDS
GSS.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 386)
Adams,M.D., Rounsley,S.D., Zhao,S., Field,C.E., Bass,S., Linher,K.,
Golden,K., Berry,K., Granger,D., Suh,E., Wible,C., Shizuya,H.,
Simon,M. and Venter,J.C.
AUTHORS
Simon M. and Venter J.C.
TITLE
Use of a random BAC End Sequence Database for Sequence-Ready Map
Building (1998)
JOURNAL
Unpublished (1998)
COMMENT
Other GSSs: CIT-HSP-2294F11.TR
Contact: Mark Adams
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: mdadams@tigr.org
Clones are available from Research Genetics (info@resgen.com). BAC
end search page:
http://www.tigr.org/tcdb/hungen/bac_end_search/bac_end_search.html
Seq primer: M13-21;
Class: BAC ends.
Location/Qualifiers
1..386
/organism="Homo sapiens"
/mol_type="genomic DNA"

/db_xref="GDB:7152159"
/db_xref="taxon:9606"
/clone="2294F11"
/sex="Male"
/cell_type="Sperm"
/note="Vector: pBelobAC11; Site_1: HindIII; Site_2:
HindIII"
ORIGIN
Query Match 6.2%; Score 233; DB 28; Length 386;
Best Local Similarity 100.0%; Pred. No. 3.6e-35;
Matches 233; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 827 AGGTGAGACACACCTTGTATTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTACACAGA 886
Db 314 AGGTGAGACACACCTTGTATTATACCCCTTCCCTTTTGGAGTTTATGCACAAGTACACAGA 255
Qy 887 TGAGTCATAGACTGATGAATAGACTGATTCGCGCATAGAGTCCCAATTCACACCTG 946
Db 254 TGAGTCATAGACTGATGAATAGACTGATTCGCGCATAGAGTCCCAATTCACACCTG 195
Qy 947 ACTCTGCTAGATCACACACTGTCTGAGGAGTTCATCTATGAGACTTTGTCTACATAA 1006
Db 194 ACTCTGCTAGATCACACACTGTCTGAGGAGTTCATCTATGAGACTTTGTCTACATAA 135
Qy 1007 CAGAGACCTTGGTTTCCACACACCCCTTTATTTTAGCTAAAGCATTTCTTTCTA 1059
Db 134 CAGAGACCTTGGTTTCCACACACCCCTTTATTTTAGCTAAAGCATTTCTTTCTA 82

RESULT 16
BF240893
LOCUS
DEFINITION
601875444F2 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4091963 5',
mRNA sequence.
ACCESSION
BF240893
VERSION
BF240893.1 GI:11154817
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 809)
NIH-MGC http://mgi.nci.nih.gov/.
AUTHORS
National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL
Unpublished (1999)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cygaps@roaill.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONETECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.lnl.gov
Plate: LLCM948 row: h column: 12
High quality sequence stop: 651.
Location/Qualifiers
1..809
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4091963"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH MGC 55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggcccttcggcc); Site_2: SfiI
(ggccattatggcc); Double-stranded cDNA was prepared from
cell line RNA. 5' and 3' adaptors were used in cloning as
follows: 5' adaptor sequence: 5'-CACGCCCATTAAGGCC-3' and
3' adaptor sequence:

5'-ATTCTAGAGCGCGCGCGCATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 5.2%; Score 194; DB 10; Length 809;
Best Local Similarity 99.5%; Pred. No. 4e-28;
Matches 364; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 718 AATGAAATTAATTTCTAATCCCTGACTAACTGAATGACCTCTTTAGGCCAAAG 777
DB 137 AATGAAATTAATTTCTAATCCCTGACTAACTGAATGACCTCTTTAGGCCAAAG 196

QY 778 AGACCTCAGATGAACCTGAAGACTGAATTTGGCCATGATAGGAGGAGTGAGACAC 837
DB 137 AGACCTCAGATGAACCTGAAGACTGAATTTGGCCATGATAGGAGGAGTGAGACAC 256

QY 838 ACCTTGTTATACCCCTTCCCTTTGGAGTTTATGACAAAGTGACAGGATGAGTATAAG 897
DB 257 ACCTTGTTATACCCCTTCCCTTTGGAGTTTATGACAAAGTGACAGGATGAGTATAAG 315

QY 898 ACTGATGAATAGACTGATGTCGCAATAGAGTCCCAATTCACCTGACTCTGGTGTA 957
DB 316 ACTGATGAATAGACTGATGTCGCAATAGAGTCCCAATTCACCTGACTCTGGTGTA 375

QY 958 GATCACACACTGTCTGAGGATTCATCTATGAGACTTTGCTACATAACAGACCTTG 1017
DB 376 GATCACACACTGTCTGAGGATTCATCTATGAGACTTTGCTACATAACAGACCTTG 435

QY 1018 GTTCCCAACCCCTTATTTAGCTTAAGACTTTCTTACTGACTCTTAACTGCTTT 1077
DB 436 GTTCCCAACCCCTTATTTAGCTTAAGACTTTCTTACTGACTCTTAACTGCTTT 495

QY 1078 AGACAA 1083
DB 496 AGACAA 501

RESULT 17
BF243193
LOCUS
DEFINITION 601877334F1 NIH_MGC_55 Homo sapiens cDNA clone IMAGE:4105665 5',
mRNA sequence.

ACCESSION BF243193
VERSION BF243193.1 GI:11157122
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 848)
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE NIH-MGC http://mgi.nci.nih.gov/
JOURNAL National Institutes of Health, Mammalian Gene Collection (MGC)
COMMENT Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC

CDNA Library Preparation: CLONETECH Laboratories, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCM984 row: c column: 10
High quality sequence stop: 546.

FEATURES
source
1..848
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

/clone="IMAGE:4105665"
/tissue_type="from acute myelogenous leukemia"
/lab_host="DH10B (T1 phage-resistant)"
/clone_lib="NIH_MGC_55"
/note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggccattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGCCATTATGCC-3' and 3' adaptor sequence:
5'-ATTCTAGAGCGCGCGCGCATG-dT(30)BN-3' (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN

Query Match 4.3%; Score 163; DB 10; Length 848;
Best Local Similarity 100.0%; Pred. No. 2.7e-22;
Matches 163; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3049 CATCAATGCTCTCTGCTGCTTCCCATAGCCACAGCAGATCTTAAATCTTTCCGAGAT 3108
DB 1 CATCAATGCTCTCTGCTGCTTCCCATAGCCACAGCAGATCTTAAATCTTTCCGAGAT 60

QY 3109 CTAGTTTTTCAGCAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTATGAGAC 3168
DB 61 CTAGTTTTTCAGCAAGCAGGATTTAAGAAATGTAATCTTATGTTGTTATGAGAC 120

QY 3169 AATAGAAATCATGCTCTATAGTCTTTTAACTGTAATTT 3211
DB 121 AATAGAAATCATGCTCTATAGTCTTTTAACTGTAATTT 163

RESULT 18
AG182348
LOCUS
DEFINITION Pan troglodytes DNA, clone: RP43-055101.TJ, genomic survey
sequence.

ACCESSION AG182348
VERSION AG182348.1 GI:16712028
KEYWORDS GSS.
SOURCE Pan troglodytes (chimpanzee)
ORGANISM Pan troglodytes
REFERENCE 1 (bases 1 to 643)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE BAC end sequences of Library RPCI-43
JOURNAL Unpublished
COMMENT BAC end sequences of Library RPCI-43

REFERENCE 2 (bases 1 to 643)
AUTHORS Fujiyama, A., Hattori, M., Toyoda, A., Taylor, T.D., Yada, T., Totoki, Y., Watanabe, H. and Sakaki, Y.
TITLE Direct Submission
JOURNAL Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical and Chemical Research (RIKEN), Genomic Sciences Center (GSC); 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
COMMENT (E-mail: chimpanzee@sc.riken.go.jp, URL: http://hgp.gsc.riken.go.jp/; Tel: 81-45-503-9111, Fax: 81-45-503-9170)
Clones are derived from the chimpanzee BAC library RPCI-43. This BAC end was generated during the Red process and may have higher chance of clone tracking errors.

PRIMERS
Sequencing: TJ
LIBRARY
Vector: pBACe3.6
R.site 1: EcoRI
R.site 2: EcoRI
Location/Qualifiers
1..643
/organism="Pan troglodytes"

/mol_type="genomic DNA"
 /db_xref="taxon:9598"
 /clone="RP43-055101.TU"
 /sex="male"
 /cell_type="lymphocytes"
 /clone_lib="RPCI-43 Chimpanzee Male BAC Library"

ORIGIN

Query Match 2.1%; Score 80; DB 29; Length 643;
 Best Local Similarity 100.0%; Pred. No. 1.7e-06;
 Matches 80; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3131 TTTAAGAAATGTAATCTATCTATGTTATGAGAACCAATAGATCATCTGCTGATATAG 3190
 Db 389 TTTAAGAAATGTAATCTATCTATGTTATGAGAACCAATAGATCATCTGCTGATATAG 448
 QY 3191 TGCTTTTACCTGTAAT 3210
 Db 449 TGCTTTTACCTGTAAT 468

RESULT 19
 BG779038
 LOCUS 712 bp mRNA linear EST 15-MAY-2001
 DEFINITION 602667021F1 NIH_MGC_60 Homo sapiens cDNA clone IMAGE:4806988 5',
 mRNA sequence.
 ACCESSION BG779038.1 GI:114049355
 VERSION BG779038
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs@email.nih.gov

Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: CLONTECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LNCM1657 row: i column: 05
 High quality sequence stop: 123.
 Location/Qualifiers
 1. 712

FEATURES

source
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4806988"
 /tissue_type="adenocarcinoma"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NIH_MGC_60"
 /note="Organ: prostate; Vector: pDNR-LIB (Clontech);
 Site 1: Sfii (ggcgctcgcc); Site 2: Sfii
 (ggccataggcc); Double-stranded cDNA was prepared from
 cell line RNA. 5' and 3' adaptors were used in cloning as
 follows: 5' adaptor sequence: 5'-CAGCGCCATATGGCC-3' and
 3' adaptor sequence:
 5'-ATTCTAGGCGGAGCGGCACATG-dt(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.5
 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.4%; Score 53; DB 12; Length 712;
 Best Local Similarity 100.0%; Pred. No. 0.19;

Matches 53; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 GAGAAACCGAGTCACTGTGAAAGATGGAAATTTATCTCTCGGAAACTCAG 53
 Db 42 GAGAAACCGAGTCACTGTGAAAGATGGAAATTTATCTCTCGGAAACTCAG 94

RESULT 20

CA301835
 LOCUS 344 bp mRNA linear EST 01-NOV-2002
 DEFINITION tao08210.y1 Hydra cDNA library Hydra magnipapillata cDNA 5', mRNA
 sequence.

ACCESSION CA301835
 VERSION CA301835.1 GI:24464894
 KEYWORDS EST.
 SOURCE Hydra magnipapillata

ORGANISM Hydra magnipapillata
 Eukaryota; Metazoa; Cnidaria; Hydrozoa; Hydrozoa; Anthomedusae;
 Hydridae; Hydra.

REFERENCE

AUTHORS Bode, H., Blumberg, B., Steele, R., Wigge, P., Gee, L., Nguyen, Q.,
 Martinez, D., Kibler, D., Hampson, S., Clifton, S., Pape, D., Marra, M.,
 Hillier, L., Martin, J., Wylie, T., Dante, M., Theising, B., Bowers, Y.,
 Gibbons, M., Ritter, E., Bennett, J., Ronko, I., Tsagareishvili, R.,
 Maguire, L., Kennedy, S., Waterston, R. and Wilson, R.
 WashU Hydra EST Project
 Unpublished (2002)

TITLE

JOURNAL

COMMENT

Contact: Hans Bode
 WashU Hydra EST Project
 Washington University School of Medicine
 4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu

Library was constructed by H. Bode and B. Blumberg DNA sequencing
 by: Washington University Genome Sequencing Center For information
 on obtaining a clone please contact: Hans Bode (hnbode@wustl.edu)

Seq primer: -40RP from Gibco

High quality sequence stop: 337.

FEATURES

source

Location/Qualifiers
 1. 344
 /organism="Hydra magnipapillata"
 /mol_type="mRNA"
 /db_xref="taxon:6085"
 /lab_host="DH5alpha"
 /clone_lib="Hydra cDNA library"
 /note="Vector: pSPORT6; Site 1: NotI; Site 2: SalI;
 Libraries prepared by Phil Wigge, Plant Molecular and
 Cellular Biology, Lab (Detlev Weigel), the Salk Institute
 for Biological Science, 10010 North Torrey Pines Road, La
 Jolla, CA 92037."

ORIGIN

Query Match 1.4%; Score 52; DB 14; Length 344;
 Best Local Similarity 100.0%; Pred. No. 0.62;
 Matches 52; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3711 TATAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 3762

Db 241 TATAATGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 292

RESULT 21

AV706915
 LOCUS 733 bp mRNA linear EST 09-OCT-2000
 DEFINITION AV706915 ADB Homo sapiens cDNA clone ADBXC08 5', mRNA sequence.
 ACCESSION AV706915
 VERSION AV706915.1 GI:10724188

KEYWORDS

SOURCE

ORGANISM

Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE
AUTHORS
  Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G.,
  Xiao, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z.,
  Xu, Z., Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S.,
  Zhong, M., Lu, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
TITLE
  Homo sapiens cDNA ADB clones
JOURNAL
  Unpublished (2000)
COMMENT
  Contact: Zeguang Han
  Chinese National Human Genome Center at Shanghai
  351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
  201203, P. R. China
  Tel: 86-21-50801919(ex.45)
  Fax: 86-21-50801922
  Email: hanzg@hgsc.sh.cn
  This clone is available at CHGC in Shanghai.
FEATURES
  source
    1..733
      Location/Qualifiers
        /organism="Homo sapiens"
        /mol_type="mRNA"
        /db_xref="taxon:9606"
        /clone="ADBBXC08"
        /tissue_type="Adrenal gland"
        /dev_stage="Adult"
        /lab_host="SOLR"
        /clone_lib="ADB"
        /note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
        XhoI"
ORIGIN
  Query Match      1.4%; Score 52; DB 9; Length 733;
  Best Local Similarity 100.0%; Pred. No. 0.29; Indels 0; Gaps 0;
  Matches 52; Conservative 0; Mismatches 0;
QY  3711 TATAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db   108 TATAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 159

RESULT 22
CA780165
LOCUS
  CA780165
DEFINITION
  MPL384.4_H01 MPL Sus scrofa cDNA clone pSPORT1 5', mRNA sequence.
ACCESSION
  CA780165
VERSION
  CA780165.1 GI:26018151
KEYWORDS
  EST.
SOURCE
  Sus scrofa (pig)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
  1 (bases 1 to 190)
  Center for Animal Functional Genomics.
  Generation of ESTs from mixed pig cDNA libraries
  Unpublished (2002)
  Contact: Steven P. Suchyta
  Center for Animal Functional Genomics, Department of Animal Science
  Michigan State University
  B215 Anthony Hall, East Lansing, MI 48824, USA
  Tel: 517 355 8443
  Fax: 517 432 9168
  Email: suchyta@msu.edu
  Single Pass sequencing. Bases called and alt-trimmed with phred
  v0.0204425.c. Vector identified by cross_match with the -minscore
  20 -minmatch 12 options.
  Seq primer: T7.
FEATURES
  source
    1..190
      Location/Qualifiers
        /organism="Sus scrofa"
        /mol_type="mRNA"
        /db_xref="taxon:9823"
        /clone="pSPORT1"
        /sex="Male and female"
        /tissue_type="pooled"
        /dev_stage="pooled"
        /lab_host="MPL"
        /clone_lib="MPL"
        /note="Organ: pooled; Vector: pSPORT1; Site_1: NotI;
        Site_2: SalI; Library made from pooled tissue from
        adipose, adrenal gland, blood leukocytes, brain,
        cartilage, eye, heart, intestine, kidney, liver, lung
        lymph nodes, mammary gland, myogenic satellite cells,
        ovary, pancreas, pituitary gland, placenta, skin, spinal
        cord, spleen, stomach, tendon, testes, uterus, and
        vascular from various developmental and physiological
        stages."
ORIGIN
  Query Match      1.4%; Score 51; DB 14; Length 322;
  Best Local Similarity 100.0%; Pred. No. 1;
  Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

ORIGIN
  Query Match      1.4%; Score 51; DB 14; Length 190;
  Best Local Similarity 100.0%; Pred. No. 1.7; Indels 0; Gaps 0;
  Matches 51; Conservative 0; Mismatches 0;
QY  3712 ATAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db   92 ATAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 142

RESULT 23
CA779205
LOCUS
  CA779205
DEFINITION
  MPL384.8_120 MPL Sus scrofa cDNA clone pSPORT1 5', mRNA sequence.
ACCESSION
  CA779205
VERSION
  CA779205.1 GI:26017100
KEYWORDS
  EST.
SOURCE
  Sus scrofa (pig)
  ORGANISM
    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
    Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE
  1 (bases 1 to 322)
  Center for Animal Functional Genomics.
  Generation of ESTs from mixed pig cDNA libraries
  Unpublished (2002)
  Contact: Steven P. Suchyta
  Center for Animal Functional Genomics, Department of Animal Science
  Michigan State University
  B215 Anthony Hall, East Lansing, MI 48824, USA
  Tel: 517 355 8443
  Fax: 517 432 9168
  Email: suchyta@msu.edu
  Single Pass sequencing. Bases called and alt-trimmed with phred
  v0.0204425.c. Vector identified by cross_match with the -minscore
  20 -minmatch 12 options.
  Seq primer: T7.
FEATURES
  source
    1..322
      Location/Qualifiers
        /organism="Sus scrofa"
        /mol_type="mRNA"
        /db_xref="taxon:9823"
        /clone="pSPORT1"
        /sex="Male and female"
        /tissue_type="pooled"
        /dev_stage="pooled"
        /lab_host="DH10B"
        /clone_lib="MPL"
        /note="Organ: pooled; Vector: pSPORT1; Site_1: NotI;
        Site_2: SalI; Library made from pooled tissue from
        adipose, adrenal gland, blood leukocytes, brain,
        cartilage, eye, heart, intestine, kidney, liver, lung
        lymph nodes, mammary gland, myogenic satellite cells,
        ovary, pancreas, pituitary gland, placenta, skin, spinal
        cord, spleen, stomach, tendon, testes, uterus, and
        vascular from various developmental and physiological
        stages."
ORIGIN
  Query Match      1.4%; Score 51; DB 14; Length 322;
  Best Local Similarity 100.0%; Pred. No. 1;
  Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 3712 ATAAATGAAAAA 3762
 LOCUS
 DEFINITION
 Db 258 ATAAATGAAAAA 308

RESULT 24
 BX512688
 LOCUS
 DEFINITION BX512688 Soares mouse 3NME12 5 Mus musculus cDNA clone
 IMAGE:5716133 5', mRNA sequence.
 ACCESSION BX512688
 VERSION BX512688.1 GI:32242549
 KEYWORDS
 SOURCE Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 502)
 COMMENT Heil, O., Ebert, L., Neubert, P., Peters, M., Radelof, U., Schneider, D.
 Mouse Unigeneset - RZPD2
 Unpublished (2003)
 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGE:5716133
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
 http://www.rzpd.de/cloncards/cgi-bin/showLib.pl.cgi?response?libNo=981 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de

FEATURES
 source
 1..502
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5716133"
 /sex="unknown"
 /tissue_type="fetus"
 /dev_stage="12.5dpc total fetus"
 /lab_host="DH10B"
 /clone_lib="Soares mouse 3NME12 5"
 /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15' TGTTACCAATCTGAAGTGGGCGCGCGCTATTTTATTTT 3', on total mouse RNA provided by Minoru Ko, Wayne State Univ.; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 1.4%; Score 51; DB 13; Length 502;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3712 ATAAATGAAAAA 3762
 Db 412 ATAAATGAAAAA 462

RESULT 25
 BX512688
 LOCUS
 DEFINITION BX512688 Soares mouse 3NME12 5 Mus musculus cDNA clone
 IMAGE:5716133 5', mRNA sequence.
 ACCESSION BX512688
 VERSION BX512688.1 GI:32242549
 KEYWORDS
 SOURCE Mus musculus (house mouse)

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 JOURNAL 1 (bases 1 to 502)
 COMMENT Heil, O., Ebert, L., Neubert, P., Peters, M., Radelof, U., Schneider, D.
 Mouse Unigeneset - RZPD2
 Unpublished (2003)
 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
 RZPD; IMAGE:5716133
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;
 Mouse Unigeneset - RZPD2 (RZPDLIB No.981)
 http://www.rzpd.de/cloncards/cgi-bin/showLib.pl.cgi?response?libNo=981 Contact: Ina Rolfs
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
 Heubnerweg 6, D-14059 Berlin, Germany
 Tel: +49 30 32639 101
 Fax: +49 30 32639 111
 www.rzpd.de

FEATURES
 source
 1..502
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="IMAGE:5716133"
 /sex="unknown"
 /tissue_type="fetus"
 /dev_stage="12.5dpc total fetus"
 /lab_host="DH10B"
 /clone_lib="Soares mouse 3NME12 5"
 /note="Organ: whole fetus; Vector: pT7T3D-Pac (Pharmacia) with a modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA was primed with a Not I - oligo(dT) primer 15' TGTTACCAATCTGAAGTGGGCGCGCGCTATTTTATTTT 3', on total mouse RNA provided by Minoru Ko, Wayne State Univ.; double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Not I and cloned into the Not I and Eco RI sites of the modified pT7T3 vector. Library went through one round of normalization, and was constructed by Bento Soares and M. Fatima Bonaldo."

ORIGIN
 Query Match 1.4%; Score 51; DB 13; Length 502;
 Best Local Similarity 100.0%; Pred. No. 0.65;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3712 ATAAATGAAAAA 3762
 Db 412 ATAAATGAAAAA 462

BQ571861/c
 LOCUS
 DEFINITION
 ACCESSION BQ571861
 VERSION BQ571861.1 GI:21475178
 KEYWORDS
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE
 AUTHORS NIH-MGC http://mgc.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgaabs-r@mail.nih.gov
 Tissue Procurement: Dr. James Lin, University of Iowa
 CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
 DNA sequencing by: Dr. M. Bento Soares, University of Iowa
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 This clone was contributed by the Brain Molecular Anatomy Project (BMAP).
 The following repetitive elements were found in this cDNA sequence: 17-74, >POLY_A#Simple_repeat (matched complement)
 Seq primer: pYX-5.

FEATURES
 source
 1..649
 /organism="Mus musculus"
 /mol_type="mRNA"
 /strain="C57BL/6"
 /db_xref="taxon:10090"
 /clone="IMAGE:5716133"
 /tissue_type="whole brain"
 /dev_stage="embryo 12.5 dpc"
 /lab_host="DH10B (T1 phage resistant)"
 /clone_lib="NIH BMAP FCO"
 /note="Organ: Brain; Vector: pYX-Asc; Site 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured mRNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with an oligo-dT primer containing a Not I site. Double stranded cDNA was size selected according to mRNA size fraction. ligated with EcoR I adaptor digested with Not I and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail, is TGAGAGAGCC. This library was created for the University of Iowa Mouse Brain Molecular Anatomy Project (BMAP). 'Gene discovery in the Developing Mouse Nervous System', supported by National Institutes of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN
 Query Match 1.4%; Score 51; DB 13; Length 649;
 Best Local Similarity 100.0%; Pred. No. 0.51;
 Matches 51; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 3712 ATAAATGAAAAA 3762
 Db 73 ATAAATGAAAAA 23

RESULT 26
 BX370381
 LOCUS
 DEFINITION BX370381 Homo sapiens HELA CELLS COT 25-NORMALIZED Homo sapiens
 cDNA clone CS0DK003IN10 5-PRIME, mRNA sequence.
 ACCESSION BX370381
 VERSION BX370381.1 GI:30447935

RESULT 25
 BX512688
 LOCUS
 DEFINITION BX512688 Soares mouse 3NME12 5 Mus musculus cDNA clone
 IMAGE:5716133 5', mRNA sequence.
 ACCESSION BX512688
 VERSION BX512688.1 GI:32242549

```

KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 931)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5501.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CSOBAG044ZE08_CS04189_1&cluster=5501.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CSOBAG044ZE08_CS04189_1.

FEATURES
source
Location/Qualifiers
1..931
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODK003VN10"
/cell_type="HELA CELLS COT 25-NORMALIZED"
/clone_lib="HELA"
/clone_lib="Homo sapiens HELA CELLS COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.4%; Score 51; DB 13; Length 931;
Best Local Similarity 100.0%; Pred. No. 0.35; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 3712 ATAAATGAAAAA... 1201 bp mRNA linear EST 08-MAY-2003
|||||
Db 154 ATAAATGAAAAA... 1201 bp mRNA linear EST 08-MAY-2003
|||||

RESULT 27
BX387156 1201 bp mRNA linear EST 08-MAY-2003
LOCUS
DEFINITION
BX387156 Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED Homo sapiens
CDNA clone CSODC010YP24 5-PRIME, mRNA sequence.
ACCESSION
BX387156 1 GI:30455496
KEYWORDS
SOURCE
ORGANISM
Homo sapiens (human)
REFERENCE
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
TITLE 1 (bases 1 to 1201)
JOURNAL Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
COMMENT Full-length cDNA libraries and normalization
Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 3161.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS5AA004ZA04QP1&cluster=3161.r. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS5AA004ZA04QP1.

```

```

FEATURES
source
Location/Qualifiers
1..1201
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODC010YP24"
/tissue_type="NEUROBLASTOMA COT 25-NORMALIZED"
/clone_lib="Homo sapiens NEUROBLASTOMA COT 25-NORMALIZED"
/notes="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.4%; Score 51; DB 13; Length 1201;
Best Local Similarity 100.0%; Pred. No. 0.27; Indels 0; Gaps 0;
Matches 51; Conservative 0; Mismatches 0;

QY 3712 ATAAATGAAAAA... 80 bp mRNA linear EST 18-AUG-2003
|||||
Db 703 ATAAATGAAAAA... 80 bp mRNA linear EST 18-AUG-2003
|||||

RESULT 28
CF328728 80 bp mRNA linear EST 18-AUG-2003
LOCUS
DEFINITION
CF328728 Rice callus plasmid cDNA library (NACL) Oryza
sativa cDNA clone NACL--03-M07, mRNA sequence.
ACCESSION
CF328728 1 GI:33905694
KEYWORDS
SOURCE
ORGANISM
Oryza sativa
REFERENCE
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erbartoideae; Oryzaceae; Oryza.
AUTHORS Kim, J.S., Jun, X.M., Cheong, P.J., Kim, M.J., Lee, T.H., Shin, Y.C.,
Song, S.I., Kim, J.K., Kim, Y.-K. and Nam, B.H.
TITLE Large-scale Sequencing Analysis of Rice ESTs
JOURNAL Unpublished (2003)
COMMENT Contact: Nahm B.H.
Genomics and Genetics Institute, GreenGene Biotech Inc.; Division
of Bioscience and Bioinformatics, Myongji University
Yongin, Gyeonggi, Korea
Tel: 82 31 330 6193
Fax: 82 31 321 6355
Email: bhnahm@gbio.com, bhnahm@bio.myongji.ac.kr.

FEATURES
source
Location/Qualifiers
1..80
/organism="Oryza sativa"
/mol_type="mRNA"
/cultivar="Nackdong"
/db_xref="taxon:4530"
/clone="NACL--03-M07"
/tissue_type="callus"
/dev_stage="proliferated callus on 2N6 media for 30 days"
/lab_host="E.coli DH10B"
/clone_lib="Rice callus plasmid cDNA library (NACL)"
/notes="Vector: PCR-TOPO; Site 1: EcoRI; mRNA was capped
with oligoribonucleotides and then used as templates for
RT-PCR."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 80;
Best Local Similarity 100.0%; Pred. No. 6.3; Indels 0; Gaps 0;
Matches 50; Conservative 0; Mismatches 0;

QY 3713 TAAATGAAAAA... 3762
|||||
Db 56 TAAATGAAAAA... 3762
|||||

```

```

RESULT 29
CF929023      142 bp  mRNA  linear  EST 12-NOV-2003
LOCUS        laf76f02.y1 SlEP Mus musculus cDNA 5', mRNA sequence.
DEFINITION   CF929023
ACCESSION   CF929023
VERSION     CF929023.1 GI:38276827
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 142)
AUTHORS      Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
COMMENT      Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .142
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitors"
/lab_host="DMSalpha Ultra Max cells (Invitrogen)"
/clone_lib="SlEP"
/notes="Vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 142;
Best Local Similarity 100.0%; Pred. No. 3.5;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 91 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 140

RESULT 30
CF755431      146 bp  mRNA  linear  EST 16-OCT-2003
LOCUS        lae60c03.y2 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
DEFINITION   mRNA sequence.
ACCESSION   CF755431
VERSION     CF755431.1 GI:37699236
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 146)
AUTHORS      Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
COMMENT      Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division

```

```

Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .146
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/lab_host="DMSalpha"
/clone_lib="Colon epithelia progenitors cDNA"
/notes="Vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 146;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 97 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 146

RESULT 31
CB645057      148 bp  mRNA  linear  EST 17-JUN-2003
LOCUS        lab18a09.y1 Normalized Small intestinal epithelia progenitors Mus
DEFINITION   musculus cDNA 5', mRNA sequence.
ACCESSION   CD645057
VERSION     CD645057.1 GI:31817557
KEYWORDS     EST.
SOURCE       Mus musculus (house mouse)
ORGANISM     Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE    1 (bases 1 to 148)
AUTHORS      Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
COMMENT      Contact: Rose Tidwell
WashU Stem cell EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Original lib:1st strand of cDNA was synthesized with reverse
transcriptase and oligo(dT) beads, then cDNA was amplified by PCR
using modified SMART primers. The final cDNA was cloned in pAMP1
vector in annealing reaction with Uracil DNA Glycosylase (UDG).
Normalized by generating biotin labeled linear PCR using original
lib cut with Sal I as template, followed with Sau3A digest, and
hybridizing with ssDNA from original lib. RNA provided by Jeff
Gordon lab and library construction by Rose Tidwell and Mike
Lovett.
Putative full length read
vector to vector length is

```

```

Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .148
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitor cells"
/lab_host="X11- Blue MRF"
/clone_lib="Normalized Small intestinal epithelia
progenitors"
/notes="vector: pAMP1; Site 1: UDG cloning; Site 2: UDG
cloning; Original lib: 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Normalized by
generating Biotin labeled linear PCR using original lib
cut with Sal I as template, followed with SauaI digest,
and hybridizing with seqDNA from original lib. RNA provided
by Jeff Gordon lab and library construction by Rose
Tidwell and Mike Lovett."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 148;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 148 bp mRNA linear EST 11-AUG-2003
lab41h07.y1 Normalized colon epithelial progenitors cDNA library
Mus musculus cDNA 5', mRNA sequence.
CF259903
VERSION EST.
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 148)
Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,
Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennet, J.,
Ronko, I., Tsagarisvilli, R., Belaygorod, L., Grow, A., Maguire, L.,
Waterston, R. and Wilson, R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
materials provided by T.S. Stappenbeck and J.I. Gordon. Library
constructed by R. Tidwell Y. Korshunova, M. Lovett. Library was
re-arrayed by R. Tidwell.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .148
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="laser-captured colonic epithelial
progenitors cells from conventionally raised C57/Bl6 mice"
/lab_host="X11 Blue MRF"
/clone_lib="Normalized colon epithelial progenitors cDNA

FEATURES
source

```

```

library"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
materials provided by T.S. Stappenbeck and J.I. Gordon.
Library constructed by R. Tidwell Y. Korshunova, M. Lovett.
Library was re-arrayed by R. Tidwell."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 148;
Best Local Similarity 100.0%; Pred. No. 3.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 149 bp mRNA linear EST 22-SEP-2003
lad99f06.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
CF545723
VERSION EST.
KEYWORDS
SOURCE
ORGANISM
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 149)
Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,
Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennet, J.,
Ronko, I., Tsagarisvilli, R., Belaygorod, L., Grow, A., Maguire, L.,
Waterston, R. and Wilson, R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .149
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/lab_host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

FEATURES
source

```


lab94b10.v1 SiEP Mus musculus cDNA 5', mRNA sequence.

KEYWORDS	SOURCE	ORGANISM	REFERENCE	AUTHORS	TITLE
EST.	Mus musculus (house mouse)				
	Mus musculus				
	Eukaryota; Metazoa;				
	Chordata; Vertebrata; Euteleostomi;				
	Mammalia; Eutheria;				
	Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
	1. (Bases 1 to 174)				
	Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,				
	Wyllie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J.,				
	Ronko, I., Tsagarishevli, R., Belaygorod, L., Grow, A., Maguire, L.,				
	Waterston, R. and Wilson, R.				
	WashU Stem Cell EST Project				

Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)

Library materials provided by: Stappendeck, M.S. and Gordon, C.R.
Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is

```

FEATURES
source
1. .174
Location/Qualifiers
seq primer: -40KP from GIBCO.

```

```

/acc_type="maxres"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitors"
/lab_host="DH5alpha Ultra Max cells (Invitrogen)"
/clone_lib="Sisp"

```

```

/clone_lib="SISF"
/notes="Vector: pAMP1. The library was synthesized with
modified SMART primers with drrp at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

```

QY	3713	3762
QY	TAANTGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA	AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA

RESULT 39
BM177915
LOCUS
 BM177915 175 bp mRNA linear
DEFINITION ssj6ta02.y1 Gm-cl072 Glycine max cDNA clone SOYBEAN CLONE ID:
 Gm-cl072-4612 5', mRNA sequence.
ACCESSION BM177915

REFERENCE 1 (bases 1 to 175)
Glycine.
rosids; eucosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Eukaryota; Viridiplantae; streptophyta; Embryophyta; core embryophyta

AUTHORS
 SNOENAKER, R., Keam, P., Voklin, L., Bepko, J., Kucaba, T., Martin, J.,
 Xianma, A., Bolla, B., Marra, M., Hallier, L., Kucaba, T., Martin, J.,
 Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
 Bowers, Y., Person, B., Swaller, T., Gibbons, M., Page, D., Harvey, N.,
 Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
 McCann, R., Waterston, R. and Wilson, R.
 Public Software, 250 Project

TITLE	Public Soybean EST Project
JOURNAL	Unpublished (1999)
COMMENT	Contact: Shoemaker R/Public Soybean EST Project

Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact: ccu@resgen.com web site:
www.resgen.com
Seq primer: -40RP from Gibco
High quality sequence stop: 169.

FEATURES

Location/Qualifiers

1. .175

/organism="Glycine max"

/mol_type="mRNA"

/db_xref="taxon:3847"

/clone="SOYBEAN CLONE ID: Gm-cl072-4612"

/tissue_type="seedlings induced for symptoms of SDS

(Sudden Death Syndrome) disease"

/dev_stage="2-3 weeks old"

/lab_host="DH10B"

/clone_lib="Gm-cl072"

/note="vector: pBluescript II SK+; Site 1: EcoRI; Site 2: XhoI; The cDNA library was constructed from mRNA isolated from 2-3 week old seedlings that were induced for symptoms of SDS (Sudden Death Syndrome) disease by the translocation of culture filtrate of Fusarium solani f. sp. glycines (Plant Cell Report 18:375-380). Cultivar PI 567374 is partially resistant to the disease SDS. Plant tissue (expanded leaves, folded leaves, and new shoots) were collected at 1, 6, 24, and 48 hrs. after inoculation and their mRNA pooled equally for cDNA construction. The library was prepared using the Stratagene pBluescript II SK(+) library construction kit. Complementary DNA was synthesized from mRNA using a primer consisting of a poly(dT) sequence with an XhoI restriction site. EcoRI adaptors were ligated to the blunt-ended cDNA fragments followed by XhoI digestion. The cDNA insert is protected from XhoI digestion via methylation during first strand synthesis. The cDNA fragments were directionally cloned into the EcoRI-XhoI restriction site of the pBluescript vector. The ligated cDNA fragments were transformed into E.coli ElectroMax DH10B host cells. Plants were inoculated by Shuxian Li (Glen Hartman lab, University of Illinois). Library was constructed by Steve Clough (Lila Vodkin lab, University of Illinois)."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 175;

Best Local Similarity 100.0%; Pred. No. 2.9;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAATGTAATGAAA 3762

DB 68 TAAATGAAA 117

RESULT 40

LOCUS

CF379915 lac74b06.y1 S1EP Mus musculus cDNA 5', mRNA linear EST 27-AUG-2003

DEFINITION

ACCESSION

CF379915

VERSION

CF379915.1 GI:34323337

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 183)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J., Ronko,I., Tsagarishvili,R., Belaygorod,L., Grow,A., Maguire,L.,

Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.

Location/Qualifiers

1. .183

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="small intestinal epithelial progenitors"

/lab_host="DHSalpha Ultra Max cells (Invitrogen)"

/clone_lib="S1EP"

/note="vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

Query Match 1.3%; Score 50; DB 14; Length 183;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

FEATURES

source

1. .183

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

/tissue_type="small intestinal epithelial progenitors"

/lab_host="DHSalpha Ultra Max cells (Invitrogen)"

/clone_lib="S1EP"

/note="vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 183;

Best Local Similarity 100.0%; Pred. No. 2.7;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAATGTAATGAAA 3762

DB 94 TAAATGAAA 143

RESULT 41

LOCUS

CF970684 lac79a01.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',

DEFINITION

ACCESSION

CF970684

VERSION

CF970684.1 GI:38496771

KEYWORDS

SOURCE

Mus musculus (house mouse)

ORGANISM

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 183)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J., Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J., Ronko,I., Tsagarishvili,R., Belaygorod,L., Grow,A., Maguire,L., Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.

Location/Qualifiers

1. .183

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

FEATURES

source

1. .183

/organism="Mus musculus"

/mol_type="mRNA"

/db_xref="taxon:10090"

```

/tissue type="colon epithelia progenitor cells"
/lab_host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match          1.3%; Score 50; DB 14; Length 183;
Best Local Similarity 100.0%; Pred. No. 2.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 208 bp mRNA linear EST 24-NOV-2003
lab14d09.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
CF970587
ACCESSION CF970587
VERSION CF970587
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 208)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarisvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 208.
Location/Qualifiers
1..208
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue type="colon epithelia progenitor cells"
/clone_lib="DH5alpha"
/lab_host="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match          1.3%; Score 50; DB 14; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 208 bp mRNA linear EST 24-NOV-2003
lab14d09.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
CF970587
ACCESSION CF970587
VERSION CF970587
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 208)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarisvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 208.
Location/Qualifiers
1..208
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue type="colon epithelia progenitor cells"
/clone_lib="DH5alpha"
/lab_host="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match          1.3%; Score 50; DB 14; Length 208;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

Qy 3713 TAAATGAAAAA 215 bp mRNA linear EST 03-SEP-2003
lab14d09.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
CF425041
ACCESSION CF425041
VERSION CF425041
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 215)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarisvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 215.
Location/Qualifiers
1..215
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue type="colon epithelia progenitor cells"
/clone_lib="DH5alpha"
/lab_host="Colon epithelia progenitors cDNA"
/notes="vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match          1.3%; Score 50; DB 14; Length 215;
Best Local Similarity 100.0%; Pred. No. 2.3;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 230 bp mRNA linear EST 20-AUG-2003
lab55c12.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
CF352096
ACCESSION CF352096
VERSION CF352096
KEYWORDS EST.
SOURCE Mus musculus (house mouse)

```


materials provided by J. Gordon lab."

```

ORIGIN
  Query Match      1.3%; Score 50; DB 14; Length 249;
  Best Local Similarity 100.0%; Pred. No. 2;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 94 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 143

RESULT 47
CF805201
LOCUS
DEFINITION      CF805201 250 bp mRNA linear EST 24-OCT-2003
IMAGE:lad64c06.y1 Gastric Epithelial Progenitor 2 Mus musculus cDNA 5',
mRNA sequence.
ACCESSION      CF805201
VERSION
KEYWORDS
SOURCE
ORGANISM      Mus musculus (house mouse)

REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 250)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarisvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
WashU Stem cell EST Project
TITLE
JOURNAL
COMMENT      Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library materials
provided by Mills JC & Gordon JI.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.

FEATURES
source
1..250
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="laser-captured isthmal cells from tox176
transgenic mice"
/dev_stages="adult"
/lab_host="DH5alpha"
/clone_lib="Gastric Epithelial Progenitor 2"
/notes="Vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library
materials provided by Mills JC & Gordon JI."

ORIGIN
  Query Match      1.3%; Score 50; DB 14; Length 250;
  Best Local Similarity 100.0%; Pred. No. 2;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 196 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 245

RESULT 48
AW022602
LOCUS
DEFINITION      AW022602 255 bp mRNA linear EST 13-SEP-1999
IMAGE:df41b12.y1 Morton Fetal Cochlea Homo sapiens cDNA clone
mRNA sequence.
ACCESSION      AW022602
VERSION
KEYWORDS
SOURCE
ORGANISM      Homo sapiens (human)

REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 255)
Robertson,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
and Morton,C.C.
Isolation of novel and known genes from a human fetal cochlear cDNA
library using subtractive hybridization and differential screening
Genomics 23, 42-50 (1994)
95130111
7829101
COMMENT      Contact: Morton, C. C.
Departments of Pathology and Obstetrics, Gynecology and
Reproductive Biology
Brigham and Women's Hospital
75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
Tel: 617 732 7980
Fax: 617 732 6996
Email: cmorton@rics.bwh.harvard.edu
DNA sequencing and analyses were performed by National Institutes
of Health Intramural Sequencing Center (NISC; see
http://www.nisc.nih.gov).
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Plate: LLAM6186 row: C column: 23
Seq primer: M13RPI reverse primer (ABI).

FEATURES
Location/Qualifiers
1..255
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2485966"
/tissue_type="cochlea"
/dev_stages="16-22 week fetus"
/clone_lib="Morton Fetal Cochlea"
/notes="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI;
Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb, Uni-ZAP
XR Vector. Library constructed by N. Robertson, C. Morton.
~5' adaptor sequence: 5' GAATTCGCGACGAG 3' ~3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN
  Query Match      1.3%; Score 50; DB 9; Length 255;
  Best Local Similarity 100.0%; Pred. No. 2;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 195 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 244

RESULT 49
BE332081
LOCUS
DEFINITION      BE332081 259 bp mRNA linear EST 14-JUL-2000
IMAGE:ut07901.y1 Soares_thymus_2NDMT Mus musculus cDNA clone
mRNA sequence.
ACCESSION      BE332081
VERSION
KEYWORDS
SOURCE
ORGANISM      Mus musculus (house mouse)

REFERENCE
AUTHORS      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 259)
Soares,N.G., Khetarpal,U., Gutierrez-Espeleta,G.A., Bieber,F.R.
and Morton,C.C.
Isolation of novel and known genes from a human fetal cochlear cDNA
library using subtractive hybridization and differential screening
Genomics 23, 42-50 (1994)
95130111
7829101
COMMENT      Contact: Morton, C. C.
Departments of Pathology and Obstetrics, Gynecology and
Reproductive Biology
Brigham and Women's Hospital
75 Francis Street, Harvard Medical School, Boston, MA 02115, USA
Tel: 617 732 7980
Fax: 617 732 6996
Email: cmorton@rics.bwh.harvard.edu
DNA sequencing and analyses were performed by National Institutes
of Health Intramural Sequencing Center (NISC; see
http://www.nisc.nih.gov).
This clone is available royalty-free through LLNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Plate: LLAM6186 row: C column: 23
Seq primer: M13RPI reverse primer (ABI).

FEATURES
Location/Qualifiers
1..259
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2485966"
/tissue_type="cochlea"
/dev_stages="16-22 week fetus"
/clone_lib="Morton Fetal Cochlea"
/notes="Organ: ear; Vector: pBluescript SK-; Site 1: EcoRI;
Site 2: XhoI; Reference: Genomics 23, 42-50 (1994) Cloned
unidirectionally. Primer: Oligo dT. Fetal cochlea, normal.
37% of inserts <0.5 kb, 56% 0.5-1.0 kb, 7% >1 kb, Uni-ZAP
XR Vector. Library constructed by N. Robertson, C. Morton.
~5' adaptor sequence: 5' GAATTCGCGACGAG 3' ~3' adaptor
sequence: 5' CTCGAGTTTTTTTTTTTTTTT 3'"

ORIGIN
  Query Match      1.3%; Score 50; DB 9; Length 255;
  Best Local Similarity 100.0%; Pred. No. 2;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 195 TAAATGAAAAA.....AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 244

```

```

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgaps@emall.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1071380
Seq primer: -40RP from Gibco
High quality sequence stop: 166.
Location/Qualifiers
1. 259
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:3327216"
/sex="male"
/tissue_type="Thymus"
/dev stage="4 weeks"
/lab host="DH10B"
/clone_lib="Soares thymus 2N5MT"
/note="vector: pT7T3D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTACCAATCTGAGTGGAGCGCGCGTGTGTGTGTGTGTGTGTGTGTGTGT
3']; double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not I
and Eco RI sites of the modified pT7T3 vector. RNA
provided by Dr. Bertrand Jordan. Library went through two
rounds of normalization, and was constructed by Bento
Soares and M.Fatima Bonaldo."

ORIGIN
Query Match 1.3%; Score 50; DB 10; Length 259;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA... 261 bp mRNA linear EST 03-DEC-2001
||||| 5' 3'
Db 186 TAAATGAAAAA... 261 bp mRNA linear EST 03-DEC-2001
||||| 5' 3'

RESULT 50
AW733655
LOCUS
DEFINITION
Gm-c1016-10692 5', mRNA sequence.
ACCESSION
AW733655
VERSION
AW733655.1 GI:7639330
KEYWORDS
EST.
SOURCE
Glycine max (soybean)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Euphorbia; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids I; Fabales; Fabaceae; Papilionoideae; Phaseoleae;
Glycine.
1 (bases 1 to 261)
Shoemaker, R., Keim, P., Vodkin, L., Erpelting, J., Coryell, V.,
Khanna, A., Bolla, B., Marra, M., Hillier, L., Kucaba, T., Martin, J.,
Beck, C., Wylie, T., Underwood, K., Steptoe, M., Theising, B., Allen, M.,
Bowers, Y., Person, B., Swaller, T., Gibbons, M., Pape, D., Harvey, N.,
Schurk, R., Ritter, E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
Public Soybean EST Project
Unpublished (1999)
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
Contact: Shoemaker R/Public Soybean EST Project
Public Soybean EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800

Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available through: ResGen, Invitrogen Corp. 2130
South Memorial Parkway Huntsville, AL 35801 For further information
call: (800)-533-4363 or contact via email: ccu@resgen.com
Insert Length: 308 Std Error: 0.00
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. 261
/organism="Glycine max"
/mol_type="mRNA"
/db_xref="taxon:3847"
/clone="GENOME SYSTEMS CLONE ID: Gm-c1016-10692"
/tissue_type="immature flowers of field grown plants"
/lab host="XL10-Gold"
/clone_lib="Gm-c1016"
/note="Vector: pBluescript II XR; Site 1: EcoRI; Site 2:
XhoI; This cDNA library was constructed from mRNA isolated
from immature flowers of field grown plants. The cDNA
library was prepared using the Stratagene pluescript II
XR library construction kit. Complementary DNA was
synthesized from mRNA using a primer consisting of a poly
(dT) sequence with a XhoI restriction site. EcoRI adaptors
were ligated to the blunt-ended cDNA fragments followed by
XhoI digestion. The cDNA fragments were directionally
cloned into the EcoRI-XhoI restriction site of the
pluescript vector. The ligated cDNA fragments were
transformed into XL10-Gold host cells. This library was
constructed by Dr. Randy Shoemaker and Dr. John
Erpelting."

ORIGIN
Query Match 1.3%; Score 50; DB 10; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA... 261 bp mRNA linear EST 17-JUN-2003
||||| 5' 3'
Db 195 TAAATGAAAAA... 261 bp mRNA linear EST 17-JUN-2003
||||| 5' 3'

RESULT 51
CD645268
LOCUS
DEFINITION
lab20h12.y1 Normalized Small intestinal epithelia progenitors Mus
musculus cDNA 5', mRNA sequence.
ACCESSION
CD645268
VERSION
CD645268.1 GI:31818114
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 261)
Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J.,
Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J.,
Ronko, I., Tsagaris, V., Belaygorod, L., Grow, A., Maguire, L.,
Waterston, R. and Wilson, R.
WashU Stem Cell EST Project
Unpublished (2002)
Contact: Rose Tidwell
WashU Stem Cell EST Project
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Original lib.1st strand of cDNA was synthesized with reverse
transcriptase and oligo(dT) beads, then cDNA was amplified by PCR
using modified SMART primers. The final cDNA was cloned in pAMF1
vector in annealing reaction with Uracil DNA Glycosylase (UDG).
Normalized by generating biotin labeled linear PCR using original
lib cut with Sal I as template, followed with Sau3A digest, and

```

```

hybridizing with ssDNA from original lib. RNA provided by Jeff
Gordon lab and library construction by Rose Tidwell and Mike
Lovett.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 261.
Location/Qualifiers
1. .261
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitor cells"
/lab_host="XLI- Blue MRF"
/clone_lib="Normalized Small Intestinal epithelia
progenitors"
/note="Vector: pAMP1; Site 1: UDG cloning; Site 2: UDG
cloning; Original lib: 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Normalized by
generating biotin labeled linear PCR using original lib
cut with Sal I as template, followed with Sau3A digest,
and hybridizing with ssDNA from original lib. RNA provided
by Jeff Gordon lab and library construction by Rose
Tidwell and Mike Lovett."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 261;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||
Db 94 TAAATGAAAAA 143

RESULT 52
AW827290
LOCUS
DEFINITION
xnl1e09.y1 NCI_CGAP_L15 Homo sapiens cDNA clone IMAGE:2693416 5',
mRNA sequence.
ACCESSION
AW827290
VERSION
AW827290.1 GI:7921064
KEYWORDS
EST.
SOURCE
Homo sapiens (human)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 265)

REFERENCE
1 (bases 1 to 265)
AUTHORS
R. Emmert-Buck, M.D., Ph.D.
TITLE
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL
Unpublished (1997)
COMMENT
Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
cDNA library Preparation: Life Technologies, Inc.
cDNA library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone Distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL, send email to:
info@image.llnl.gov
Putative full length read
The vector to vector length is 337
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .265
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

```

```

/clone="IMAGE:2693416"
/tissue_type="hepatic adenoma"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_L15"
/note="Organ: liver; Vector: pCMV-SPORT4; Site 1: SalI;
Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 0.8 kb."

ORIGIN
Query Match 1.3%; Score 50; DB 10; Length 265;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||
Db 187 TAAATGAAAAA 236

RESULT 53
CF352983
LOCUS
DEFINITION
lab60b08.y1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
ACCESSION
CF352983
VERSION
CF352983.1 GI:33957496
KEYWORDS
EST.
SOURCE
Mus musculus (house mouse)
ORGANISM
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
1 (bases 1 to 269)

REFERENCE
1 (bases 1 to 269)
AUTHORS
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennett,J.,
Ronko,I., Tsagareishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
COMMENT
Contact: Jeff Gordon and Mike Lovett
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 269.
Location/Qualifiers
1. .269
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/lab_host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/note="Vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y.Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 269;
Best Local Similarity 100.0%; Pred. No. 1.9;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||

```



```

ORIGIN
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

Query Match      1.3%; Score 50; DB 14; Length 294;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 225 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 274

RESULT 57
CF384267
LOCUS
DEFINITION lac04h02.y1 SlEP Mus musculus cDNA 5', mRNA sequence.
ACCESSION CF384267.1 GI:34327691
VERSION
KEYWORDS
SOURCE Mus musculus (house mouse)
ORGANISM
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarishevili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova Y. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1..301
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitors"
/lab_host="DH5alpha Ultra Max cells (invitrogen)"
/clone_lib="SlEP"
/note="Vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

ORIGIN

Query Match      1.3%; Score 50; DB 14; Length 301;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 228 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 277

RESULT 58
AL045620
LOCUS
DEFINITION DXF2p4340045.r1 434 (synonym: htes3) Homo sapiens cDNA clone
ACCESSION AL045620
VERSION AL045620.2 GI:5936045
KEYWORDS
EUT.

```

```

SOURCE
ORIGIN Homo sapiens (human)
Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 305)
Duesterhoeft,A., Lauber,J., Mewes,H.W., Gassenhuber,J. and
Wiemann,S.
EST (Duesterhoeft, et al.)
Unpublished (1999)
On Jul 9, 1999 this sequence version replaced gi:5433743.
Contact: MIPS
MIPS
Ingolstaedter Landstr 1, D-85764 Neuberg, Germany
This is the 5' sequence of the clone insert
Clone from S. Wiemann, Molecular Genome Analysis, German Cancer
Research Center (DKFZ); Email s.wiemann@dkfz-heidelberg.de;
sequenced by Qiagen (Hilden/Germany) within the cDNA sequencing
consortium of the German Genome Project.
No sl sequence available.
This clone (DKFZp4340045) is available at the RZPD in Berlin.
Please contact the RZPD: Ressourcenzentrum, Heubnerweg 6, 14059
Berlin-Charlottenburg, GERMANY; Email: clone@rzpd.de.
Location/Qualifiers
1..305
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="DKFZp4340045"
/tissue_type="testis"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="434 (synonym: htes3)"
/note="Vector: pSport1; Site_1: NotI; Site_2: Sali"

ORIGIN

Query Match      1.3%; Score 50; DB 9; Length 305;
Best Local Similarity 100.0%; Pred. No. 1.7;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
      |||||||
Db 184 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 233

RESULT 59
CF754869
LOCUS
DEFINITION lae07d02.y2 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
ACCESSION CF754869.1 GI:37698672
VERSION
KEYWORDS Mus musculus (house mouse)
SOURCE
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 314)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarishevili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read

```

vector to vector length is
 Seq primer: -40RP from Gibco.
 Location/Qualifiers
 source
 1..314
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /tissue_type="colon epithelia progenitor cells"
 /lab_host="DH5alpha"
 /clone_lib="Colon epithelia progenitors cDNA"
 /note="Vector: pAMP1; 1st strand of cDNA was synthesized with reverse transcriptase and oligo(dT) beads, then cDNA was amplified by PCR using modified SMART primers. The final cDNA was cloned in pAMP1 vector in annealing reaction with Uracil DNA Glycosylase (UDG). Library constructed by Y. Korshunova and M. Lovett. Library materials provided by J. Gordon lab."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 314;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 321 bp mRNA linear EST 02-JUN-2003
 ||||| 3762
 Db 219 TAAATGAAAAA 268

RESULT 60
 CD422638
 LOCUS
 DEFINITION laa83g11.y1 SlEP Mus musculus cDNA 5', mRNA sequence.
 ACCESSION CD422638
 VERSION CD422638.1 GI:31328901
 KEYWORDS EST.
 SOURCE Mus musculus (house mouse)
 ORGANISM
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
 1 (bases 1 to 321)
 Tidwell, R., Clifton, S., Marra, M., Hillier, L., Pape, D., Martin, J., Wylie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Ronko, I., Teagareishvili, R., Belaygorod, L., Grow, A., Maguire, L., Waterston, R. and Wilson, R.
 WashU Stem cell EST Project
 Unpublished (2002)
 Contact: Rose Tidwell
 WashU Stem cell EST Project
 Washington University School of Medicine
 444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
 Tel: 314 286 1800
 Fax: 314 286 1810
 Email: est@watson.wustl.edu
 Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
 Library constructed by: Korshunova Y. and Lovett M. DNA sequencing by: Washington University Genome Sequencing Center For information on obtaining a clone please contact: Rose Tidwell (rtidwell@genetics.wustl.edu)
 Putative full length read
 vector to vector length is
 Seq primer: -40RP from Gibco
 High quality sequence stop: 321.
 Location/Qualifiers
 source
 1..321
 /organism="Mus musculus"
 /mol_type="mRNA"
 /db_xref="taxon:10090"
 /tissue_type="small intestinal epithelial progenitors"
 /lab_host="DH5alpha Ultra Max cells (Invitrogen)"
 /clone_lib="SlEP"
 /note="Vector: pAMP1; The library was synthesized with modified SMART primers with dUTP at the end. After treatment with UDG the cDNA was cloned in pAMP1 vector by

annealing."

ORIGIN
 Query Match 1.3%; Score 50; DB 14; Length 321;
 Best Local Similarity 100.0%; Pred. No. 1.6;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 329 bp mRNA linear EST 28-APR-1999
 ||||| 3762
 Db 219 TAAATGAAAAA 268

RESULT 61
 AU034841
 LOCUS
 DEFINITION Dictyostelium discoideum SL (H.Urushihara) Dictyostelium discoideum cDNA clone SLE570, mRNA sequence.

ACCESSION AU034841
 VERSION AU034841.1 GI:3800265
 KEYWORDS EST.
 SOURCE Dictyostelium discoideum
 ORGANISM Dictyostelium discoideum
 Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
 1 (bases 1 to 329)
 Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M., Yoshino, R., Mitra, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H., Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.
 The Dictyostelium developmental cDNA project: generation and analysis of expressed sequence tags from the first-finger stage of development
 DNA Res. 5 (6), 335-340 (1998)
 99156227
 10048482
 Contact: Hideko Urushihara
 Institute of Biological Sciences
 University of Tsukuba
 1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan
 Tel: 81-298-53-4664
 Fax: 81-298-53-6614
 Email: hideko@biol.tsukuba.ac.jp
 PROJECT = 'Dictyostelium discoideum cDNA project in Japan'
 POLYA=No.

TITLE
 JOURNAL
 MEDLINE
 PUBMED
 COMMENT
 FEATURES
 source
 1..329
 /organism="Dictyostelium discoideum"
 /mol_type="mRNA"
 /strain="AX4"
 /db_xref="taxon:44689"
 /clone="SLE570"
 /dev_stage="slug"
 /clone_lib="Dictyostelium discoideum SL (H.Urushihara)"

FEATURES

ORIGIN
 Query Match 1.3%; Score 50; DB 9; Length 329;
 Best Local Similarity 100.0%; Pred. No. 1.5;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
 ||||| 3762
 Db 273 TAAATGAAAAA 322

FEATURES

source
 RESULT 62
 BU767819
 LOCUS
 DEFINITION BU767819 SJE Schistosoma japonicum cDNA, mRNA sequence.
 ACCESSION BU767819
 VERSION BU767819.1 GI:28335135
 KEYWORDS EST.
 SOURCE Schistosoma japonicum
 ORGANISM Schistosoma japonicum
 Eukaryota; Metazoa; Platyhelminthes; Trematoda; Digenea; Strigeidida; Schistosomatoidea; Schistosomatidae; Schistosoma.

```

REFERENCE
AUTHORS
Hu.W., Yan.Q., Shen.D.K., Liu.F., Zhu.Z.D., Song.H.D., Xu.X.R.,
Wang,Z.J., Rong.Y.P., Zeng.L.C., Wu.J., Zhang.X., Wang.J.J.,
Xu.X.N., Wang.S.Y., Fu.G., Zhang.X.L., Wang.Z.Q., Brindley.P.J.,
McVane,D.P., Xue.C.L., Peng.Z., Chen.Z. and Han.Z.G.
TITLE
Evolutionary and biomedical implications of a Schistosoma japonicum
complementary DNA resource
JOURNAL
Nat. Genet. 35 (2), 139-147 (2003)
MEDLINE
22879925
COMMENT
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzg@chgc.sh.cn.

FEATURES
source
1. .336
/organism="Schistosoma japonicum"
/mol_type="mRNA"
/db_xref="taxon:6182"
/tissue_type="Whole egg"
/dev_stage="egg"
/lab_host="rabbits"
/clone_lib="SJE"

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 336;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 346 bp mRNA linear EST 27-AUG-2003
|||||
DB 270 TAAATGAAAAA 319

RESULT 63
CF382593
LOCUS
DEFINITION
lacs8907.y1 SLEP Mus musculus CDNA 5', mRNA sequence.
ACCESSION
CF382593
VERSION
EST.
KEYWORDS
Mus musculus (house mouse)
SOURCE
Mus musculus
ORGANISM
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 346)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyllie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagareishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
WashU Stem cell EST Project
WashU Plasmid Library
Contact: Jeff Gordon and Mike Lovett
Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .346
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitors"
/lab_host="DH5alpha Ultra Max cells (Invitrogen)"
/clone_lib="SIEP"

us-10-005-907-1.oligo.rst

```

```

/note="Vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 346;
Best Local Similarity 100.0%; Pred. No. 1.5;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
|||||
DB 225 TAAATGAAAAA 274

RESULT 64
BQ450975
LOCUS
DEFINITION
PFEST0ab02c10.y1 Plasmodium falciparum 3D7 asexual cDNA Plasmodium
falciparum 3D7 CDNA 5', mRNA sequence.
ACCESSION
BQ450975
VERSION
BQ450975.1 GI:21254087
KEYWORDS
EST.
SOURCE
Plasmodium falciparum 3D7
ORGANISM
Plasmodium falciparum 3D7
REFERENCE
1 (bases 1 to 350)
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
AUTHORS
Tang,K., Cole,R., Chakrabarti,D., Haywood,R., Clifton,S., Pape,D.,
Marra,M., Hillier,L., Martin,J., Wyllie,T., Dante,M., Theising,B.,
Bowers,Y., Gibbons,M., Ritter,E., Bennet,J., Jentes,E., Ronko,I.,
Tsagareishvili,R., Belaygorod,L., Franklin,C., Carr,L., Grow,A.,
Maguire,L., Richey,J., Wadkins,J., Kennedy,S., Levinso,D.,
Waterston,R., Wilson,R. and Sibley,D.
WashU Plasmodium EST Project
WashU Plasmid Library
Contact: L. David Sibley
Washington University School
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
Library was constructed by Debopam Chakrabarti DNA sequencing by:
Washington University Genome Sequencing Center For information on
obtaining a clone please contact: L. David Sibley
(sibley@borcim.wustl.edu), Washington University
Seq primer: -40UP from Gibco.
Location/Qualifiers
1. .350
/organism="Plasmodium falciparum 3D7"
/mol_type="mRNA"
/db_xref="taxon:36329"
/lab_host="DH10B (GeneHog, Invitrogen, Inc.)"
/clone_lib="Plasmodium falciparum 3D7 asexual cDNA"
/note="Vector: pBluescript SK plus; Site 1: EcoRI; Site 2:
XhoI; Library was constructed by Debopam Chakrabarti.
Total RNA samples were isolated from mixed stage
sporozoites (0.1%)-lysed P. falciparum 3D7 infected
erythrocytes by the acidic guanidinium-phenol chloroform
method. The poly A+ RNA was isolated by the polyAT-Tract
mRNA isolation system (Promega, WI) using streptavidin
MagneSphere particles. Directional cDNA libraries were
constructed by oligo(dT) priming of poly(A)+ RNA (5mg)
into EcoRI and XhoI sites of 1 ZapII vector using the Zap
cDNA synthesis kit (Stratagene, CA). The average size of the
cDNA inserts in the library was between 1.0 and 1.5kb.
Clones were mass excised using the ExAssist helper phage
(Stratagene), the phagemids were precipitated with PEG
8000 and extracted with phenol/chloroform. Phagemid DNA
was electroporated into DH10B cells."

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 350;

```

```

Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 296 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 345

RESULT 65
BE048163 354 bp mRNA linear EST 08-JUN-2000
LOCUS c246905.y1 NCI_CGAP_Brn52 Homo sapiens cDNA clone IMAGE:2291672 5',
DEFINITION mRNA sequence.
ACCESSION BE048163
VERSION BE048163.1 GI:8365141
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 354)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Christopher A. Moskaluk, M.D., Ph.D., Michael
R. Emmert-Buck, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LINL, send email to:
info@image.lnl.gov
Seq primer: -40RP from Gibco
High quality sequence stop: 297.
FEATURES
source
1..354
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2291672"
/tissue_type="tumor, 5 pooled (see description)"
/lab_host="DH10B"
/clone_lib="NCI_CGAP_Brn52"
/notes="Organ: Brain; Vector: pCMV-SPORT6; Site 1: SalI;
Site 2: NotI; This library represents the normalized
version of NCI_CGAP_Brn35. Cloned unidirectionally.
Primer: Oligo dT. Average insert size 1.19 kb. Tumor
types include: meningioma, oligodendroglioma, astrocytoma
(grade II), medulloblastoma, astrocytoma (grade IV).
Constructed by Life Technologies."
ORIGIN
Query Match 1.3%; Score 50; DB 10; Length 354;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 215 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 264

RESULT 66
BW306217 366 bp mRNA linear EST 11-NOV-2002
LOCUS BW306217 Nori Satoh unpublished cDNA library, heart Ciona
DEFINITION intestinalis cDNA clone cinfo02005 5', mRNA sequence.
ACCESSION BW306217
VERSION BW306217.1 GI:24886828
KEYWORDS EST.
SOURCE Ciona intestinalis

```

```

Ciona intestinalis
Eukaryota; Metazoa; Chordata; Urochordata; Ascidiacea; Enterogona;
Phlebobranchia; Cionidae; Ciona.
REFERENCE 1 (bases 1 to 366)
AUTHORS Satou, Y., Shin-I, T., Kohara, Y. and Satoh, N.
TITLE Expressed genes in Ciona intestinalis (2002c)
JOURNAL Unpublished (2002)
COMMENT Contact: Nori Satoh
Department of Zoology
Kyoto University
Sakyo-Ku, Kyoto, Kyoto 606-8502, Japan
Tel: 81-75-753-4081
Fax: 81-75-705-1113
Email: satoh@ascidian.zool.kyoto-u.ac.jp.
FEATURES
source
1..366
/organism="Ciona intestinalis"
/mol_type="mRNA"
/db_xref="taxon:7719"
/clone="cinfo02005"
/tissue_type="heart"
/clone_lib="Nori Satoh unpublished cDNA library, heart"
ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 366;
Best Local Similarity 100.0%; Pred. No. 1.4;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 305 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 354

RESULT 67
BG955855 387 bp mRNA linear EST 05-JUN-2001
LOCUS 359509 MARC 1PIG Sus scrofa cDNA 5', mRNA sequence.
DEFINITION BG955855
ACCESSION BG955855
VERSION BG955855.1 GI:14306096
KEYWORDS EST.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
REFERENCE 1 (bases 1 to 387)
AUTHORS Fahrnkug, S.C., Smith, T.P.L., Freking, B.A., Cho, J., White, J.,
Vallet, J., Wise, T., Rohrer, G.A., Pertea, G., Sultana, R.,
Quackenbush, J. and Keele, J.W.
TITLE Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
JOURNAL Mamm. Genome 13 (8), 475-478 (2002)
MEDLINE 22213789
PUBMED 12226715
COMMENT Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and alt trimmed with phred
v0.980904.e. Vector identified by cross_match with the -minscore 18
and -mismatch 12 options.
PCR Primers
FORWARD: AGGAAACAGCTATGACCAT
BACKWARD: GTTTCCAGTCACGCG
Plate: 124 row: 1 column: 6
Seq primer: ATTTAGGTGACACTATAG.
FEATURES
source
1..387
/organism="Sus scrofa"
/mol_type="mRNA"
/db_xref="taxon:9823"
/tissue_type="pooled"

```

/lab host="DH10B"

/clone_lib="MARC 1P1G"
/note="Vector: pCMV SPOR6; Site 1: NotI; Site 2: SalI;
Library made from pooled tissue from day 11, 13, 15, 20,
and 30 embryos."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 387;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAA 3762
Db 186 TAAATGAAAAA 235

RESULT 68

CF979559 395 bp mRNA linear EST 25-NOV-2003
LOCUS ig78c04.y1 Meloidogyne hapla female SL1 pGEM Meloidogyne hapla cDNA
DEFINITION 5', mRNA sequence.

ACCESSION CF979559
VERSION CF979559.1 GI:38512608
KEYWORDS EST.

SOURCE

ORGANISM Meloidogyne hapla
Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
Tylenchoidea; Heterodridae; Meloidogyninae; Meloidogyne.

REFERENCE

AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
Wylie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
Tsagarisvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCann, R., Waterston, R. and Wilson, R.
The Washington Univ. Nematode EST Project, 1999
Unpublished (1999)

TITLE

JOURNAL

COMMENT

Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: estewatson.wustl.edu

Cloned unidirectionally. Poly (A) + RNA was concentrated and
purified using Dynabeads (Dyna) and mRNA eluted for 1st strand
synthesis. 1st strand cDNA created using MMLV RT(PowerScript,
Clontech) and primed with oligo(dT) with XhoI site (primer
CDSIII/3-XhoI) and 5' SMART anchor added using chimeric DNA-RNA
oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR
using the SMART template was also desired. 15 PCR cycles were done
using 1st strand and primers specific to SL1 leader sequence
(SL1-NotI) and 3' end (XhoI-No-dT). cDNA was digested using
XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and
ligated to digested pGEM -11zf(+) plasmid. Chemically competent
cells were used as host. This library was provided by David Bird's
lab (david.bird@ncsu.edu), of North Carolina State University.
Putative full length read
The vector to vector length is 396

Seq primer: Sp6.

FEATURES

source

1. .395
Location/Qualifiers
/organism="Meloidogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/sex="female"
/tissue_type="whole organism"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Meloidogyne hapla female SL1 pGEM"
/note="Vector: pGEM-11zf(+) (Promega); Site 1: XhoI;
Site 2: NotI; cloned unidirectionally. Poly (A) + RNA was

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 395;
Best Local Similarity 100.0%; Pred. No. 1.3;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

Db 315 TAAATGAAAAA 364

RESULT 69

AA493109 407 bp mRNA linear EST 19-AUG-1997
LOCUS ng96e02.s1 NCI CGAP Thy1 Homo sapiens cDNA clone IMAGE:942650
DEFINITION similar to contains element L1 repetitive element ;, mRNA sequence.

ACCESSION AA493109
VERSION AA493109.1 GI:2222950
KEYWORDS EST.

SOURCE

ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: cgabs-r@mail.nih.gov
Tissue Procurement: L. Jeffrey Medeiros, M.D., Michael R.
Emmert-Buck, M.D., Ph.D.
cDNA Library Preparation: David B. Krizman, Ph.D.
cDNA Library Arrayed by: Greg Lennon, Ph.D.
DNA Sequencing by: Washington University Genome Sequencing Center
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Insert length: 601 Std Error: 0.00
Seq primer: -40m13 fwd. ET from Amersham
High quality sequence stop: 349.

FEATURES

source

1. .407
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:942650"
/tissue_type="thyroid"
/lab_host="DH10B"
/clone_lib="NCI CGAP Thy1"
/note="Vector: pAMP10; mRNA made from invasive thyroid
tumor, cDNA made by oligo-dT priming. Non-directionally
cloned. Size-selected on agarose gel, average insert size
600 bp. Reference: Krizman et al. (1996) Cancer Research
56:5380-5383."

ORIGIN

Query Match 1.3%; Score 50; DB 9; Length 407;

concentrated and purified using Dynabeads (Dyna) and mRNA
eluted for 1st strand synthesis. 1st strand cDNA created
using MMLV RT(PowerScript, Clontech) and primed with
oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5'
SMART anchor added using chimeric DNA-RNA oligo
(SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a
PCR using the SMART template was also desired. 15 PCR
cycles were done using 1st strand and primers specific to
SL1 leader sequence (SL1-NotI) and 3' end (XhoI-No-dT).
cDNA was digested using XhoI/NotI, fractionated on
Chroma-spin 400 columns (Clontech) and ligated to digested
pGEM -11zf(+) plasmid. Chemically competent cells were
used as host. This library was provided by David Bird's
lab (david.bird@ncsu.edu), of North Carolina State
University."

Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762
Db 346 TAAATGAAA 395

RESULT 70
A1133559
LOCUS
DEFINITION HA2177 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
VERSION A1133559.1 GI:6360875
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS 1 (bases 1 to 427)
Yu, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.
TITLE Expression profile analysis of a human fetal liver cDNA library
JOURNAL Unpublished (1998)
COMMENT Department of Hematology
Contact: Yongtao Yu
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: Yyt48@yahoo.com.

FEATURES
source
1. 427
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="WC1061/P3"
/clone_lib="Human fetal liver cDNA library"
/note="Vector: pCDNA1"

ORIGIN
Query Match 1.3%; Score 50; DB 9; Length 427;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762
Db 316 TAAATGAAA 365

RESULT 71
A1133559
LOCUS
DEFINITION lae68d10.v1 Colon epithelia progenitors cDNA Mus musculus cDNA 5',
mRNA sequence.
VERSION CF621874
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
AUTHORS 1 (bases 1 to 440)
Wyllie, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J.,
Ronko, I., Tsagareishvili, R., Belaygorod, L., Grow, A., Maguire, L.,
Waterston, R. and Wilson, R.
TITLE WashU Stem cell EST Project
JOURNAL Unpublished (2002)
COMMENT Contact: Jeff Gordon and Mike Lovett

WashU, Human Genetics Division
Washington University School of Medicine
1st strand of cDNA was synthesized with reverse transcriptase and
oligo(dT) beads, then cDNA was amplified by PCR using modified
SMART primers. The final cDNA was cloned in pAMP1 vector in
annealing reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library materials
provided by J. Gordon lab.
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco
High quality sequence stop: 440.

FEATURES
source
1. 440
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="colon epithelia progenitor cells"
/lab_host="DH5alpha"
/clone_lib="Colon epithelia progenitors cDNA"
/note="Vector: pAMP1; 1st strand of cDNA was synthesized
with reverse transcriptase and oligo(dT) beads, then cDNA
was amplified by PCR using modified SMART primers. The
final cDNA was cloned in pAMP1 vector in annealing
reaction with Uracil DNA Glycosylase (UDG). Library
constructed by Y. Korshunova and M. Lovett. Library
materials provided by J. Gordon lab."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 440;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAA 3762
Db 222 TAAATGAAA 271

RESULT 72
CF979974
LOCUS
DEFINITION X989e02.v1 Meloidogyne hapla female SL1 pGM Meloidogyne hapla cDNA
5', mRNA sequence.
ACCESSION CF979974.1 GI:38513023
VERSION CF979974
KEYWORDS EST.
SOURCE Meloidogyne hapla
ORGANISM Meloidogyne hapla
REFERENCE 1 (bases 1 to 441)
AUTHORS McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
Wyllie, T., Dante, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
Tsagareishvili, R., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
McCam, R., Waterston, R. and Wilson, R.
TITLE The Washington Univ. Nematode EST Project, 1999
JOURNAL Unpublished (1999)
COMMENT Contact: McCarter JP
The Washington Univ. Nematode EST Project, 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@wustl.wustl.edu
Cloned unidirectionally. Poly (A) + RNA was concentrated and
purified using Dynabeads (Dyna) and mRNA eluted for 1st strand
synthesis. 1st strand cDNA created using MMLV RT (PowerScript,
Clontech) and primed with oligo(dT) with XhoI site (primer
CDSIII/3-XhoI) and 5' SMART anchor added using chimeric DNA-RNA
Oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR

using the SMART template was also desired. 15 PCR cycles were done using 1st strand and primers specific to SL1 leader sequence (SL1-NotI) and 3' end (XhoI-No-dT). dscDNA was digested using XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent cells were used as host. This library was provided by David Bird's lab (david.bird@ncsu.edu), of North Carolina State University. Seq primer: Sp6

High quality sequence stop: 441.

FEATURES

Location/Qualifiers

1. .441
/organism="Meloiodogyne hapla"
/mol_type="mRNA"
/db_xref="taxon:6305"
/sex="female"
/tissue_type="whole organism"
/dev_stage="adult"
/lab_host="DH10B"
/clone_lib="Meloiodogyne hapla female SL1 pGEM"
/note="Vector: pGEM-11zf(+) (Promega); Site 1: XhoI;
Site 2: NotI; Cloned unidirectionally. Poly (A) + RNA was concentrated and purified using Dynabeads (Dyna) and mRNA eluted for 1st strand synthesis. 1st strand cDNA created using MMV RT(PowerScript, Clontech) and primed with oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5' SMART anchor added using chimeric DNA-RNA oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR using the SMART template was also desired. 15 PCR cycles were done using 1st strand and primers specific to SL1 leader sequence (SL1-NotI) and 3' end (XhoI-No-dT). dscDNA was digested using XhoI/NotI and ligated to digested Chroma-spin 400 columns (Clontech) and ligated to digested pGEM-11zf(+) plasmid. Chemically competent cells were used as host. This library was provided by David Bird's lab (david.bird@ncsu.edu), of North Carolina State University."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 441;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

Db 370 TAAATGAAAAA 419

RESULT 73

BM958905
LOCUS 443 bp mRNA linear EST 18-MAR-2002
DEFINITION PLATE_11_C09_05_A01 GS Lambda-Triplex, 10 day germinating spore
Library Glomus intraradices cDNA, mRNA sequence.

ACCESSION BM958905

VERSION BM958905.1 GI:19542335

KEYWORDS EST.

SOURCE Glomus intraradices

ORGANISM Glomus intraradices

REFERENCE Eukaryota; Fungi; Glomeromycota; Glomeromycetes; Glomerales;

Glomeraceae; Glomus.

1 (bases 1 to 443)

Authors Bagó, B., Zipfel, W., Williams, R.M., Jun, J., Arreola, R.,

Lammers, P.J., Pfeffer, P.E. and Shachar-Hill, Y.

Translocation and utilization of fungal storage lipid in the

arbuscular mycorrhizal symbiosis

Plant Physiol. 128 (1), 108-124 (2002)

JOURNAL 21648947

MEDLINE 11788757

COMMENT

Contact: Peter Lammers

New Mexico State University

Box 30001, MSC 3MLS Horseshoe Drive Las Cruces, NM 88003, USA

Tel: 505-646-3918

Fax: 505-646-6846

Email: plammers@ncsu.edu.

FEATURES

source

Location/Qualifiers
1. .443
/organism="Glomus intraradices"
/mol_type="mRNA"
/db_xref="taxon:4876"
/tissue_type="germinating spore tissue"
/dev_stage="10 days after germination"
/lab_host="E. coli BM 25.8"
/clone_lib="GS Lambda-Triplex, 10 day germinating spore library"
/note="Vector: Lambda Triplex2; Site 1: SfiI; Site 2: SfiIIB; Spores were germinated in the presence of 1% CO₂. At 10 days, tissue was ground in liquid nitrogen with sand and mRNA isolated with a modified hot phenol/SDS method. The SMART-PCR method (Clontech) was used for cDNA synthesis from 1 ug of total RNA followed by directional cloning in SfiI digested Triplex2 vector. Plasmid subclones in pTriplex were recovered by cre-lox excision in E. coli strain BM25.8 and sequenced from the 5' end with the 5'Triplex sequencing primer (Tpx) and the 3' end with the T7 primer (T7). (sfiI= bp in Comment line of EST file is estimated insert size)"

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 443;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762

Db 361 TAAATGAAAAA 410

RESULT 74

C91260 457 bp mRNA linear EST 20-APR-1998
LOCUS C91260 Dictyostelium discoideum SS (H.Urushihara) Dictyostelium
DEFINITION discoideum cDNA clone SSU516, mRNA sequence.

ACCESSION C91260

VERSION C91260.1 GI:3060626

KEYWORDS EST.

SOURCE Dictyostelium discoideum

ORGANISM Dictyostelium discoideum

REFERENCE Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.

1 (bases 1 to 457)

Authors Morio, T., Urushihara, H., Saito, T., Ugawa, Y., Mizuno, H., Yoshida, M.,

Yoshino, R., Mitra, B.N., Pi, M., Sato, T., Takemoto, K., Yasukawa, H.,

Williams, J., Maeda, M., Takeuchi, I., Ochiai, H. and Tanaka, Y.

Developmental cDNA in Dictyostelium discoideum

Unpublished (1998)

Contact: Hideko Urushihara

Institute of Biological Sciences

University of Tsukuba

1-1-1 Tennoudai, Tsukuba, Ibaraki 305-8572, Japan

Tel: 81-298-53-4664

Fax: 81-298-53-6614

Email: hideko@biol.tsukuba.ac.jp.

Location/Qualifiers

source

1. .457

/organism="Dictyostelium discoideum"

/mol_type="mRNA"

/strain="AX4"

/db_xref="taxon:44689"

/clone="SSJ516"

/dev_stage="slug"

/clone_lib="Dictyostelium discoideum SS (H.Urushihara)"

ORIGIN

Query Match 1.3%; Score 50; DB 13; Length 457;
Best Local Similarity 100.0%; Pred. No. 1.1;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3712 ATAAATGAAAAA 3761

```

ORIGIN

Query Match      1.3%; Score 50; DB 9; Length 464;
Best Local Similarity 100.0%; Pred.No.1.;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
        |||
Db       353 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 402

RESULT 76
LOCUS   CF546023
DEFINITION lag89a06.y1 Colon epithelia progenitors cDNA Mus musculus cdna 5', linear EST 22-SEP-2003
MUSCULUS sequence.
ACCESSION CF546023
VERSION   CF546023.1 GI:34893453
KEYWORDS EST.
SOURCE   Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Scaurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 464)
Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wyde,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Runko,I., Tsagarishevili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson.R.
WashU Stem cell EST Project
Unpublished (2002)
Contact: Jeff Gordon and Mike Lovett

TITLE
JOURNAL
COMMENT
```


/clone="IMAGE:6846599"
 /tissue type="whole brain"
 /dev stage="1, 5, and 15 days newborn"
 /lab host="DH10B (T1 phage resistant)"
 /clone lib="NTH_BMAP_GKO"
 /note="Organ: Brain; Vector: pYX-Asc; Site: 1: EcoR I; Site 2: Not I; The library was constructed according to Bonaldo, Lennon and Soares, Genome Research, 6:791-806, 1996. Denatured RNA was size fractionated on a 1% agarose gel. First strand cDNA synthesis was primed with oligo-dT primer containing a Not I site. Double strand cDNA was size selected according to mRNA size fraction, ligated with EcoR I adaptor, digested with NotI and then cloned directionally into pYX-Asc vector. The library tag sequence located between the Not I site and the polyA tail is CGAATGTAAT. This library was created for the University of Iowa Brain Anatomy Project (BMAP): 'Gene Discovery in the Developing Mouse Nervous System', supported by National Institute of Mental Health (NIMH), Hemin Chin, Ph.D., program coordinator."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 477;
 Best Local Similarity 100.0%; Pred. No. 1.1;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 410 TAAATGAAAAA 459

RESULT 78

LOCUS B1376676 483 bp mRNA linear EST 26-AUG-2003
 DEFINITION BFG3-000479 Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMPG498) Branchiostoma floridae cDNA clone MPMPG498B0638 5', mRNA sequence.

ACCESSION B1376676
 VERSION B1376676.1 GI:30911635

KEYWORDS EST.
 SOURCE Branchiostoma floridae (Florida lancelet)
 ORGANISM Branchiostoma floridae

REFERENCE 1 (bases 1 to 483)
 AUTHORS Panopoulou, G., Hennig, S., Groth, D., Krause, A., Poustka, A. J., Herwig, R., Vingron, M. and Lehrach, H.

TITLE New evidence for genome-wide duplications at the origin of vertebrates using an amphioxus gene set and completed animal genomes

JOURNAL Genome Res. 13 (6A), 1056-1066 (2003)
 MEDLINE 22683279
 PUBMED 12799346

COMMENT Contact: Panopoulou G

Laboratory 145, dept. Lehrach
 Max-Planck-Institut fuer Molekulare Genetik

Innestr. 63-73, D-14195 Berlin, Germany

Tel: +49 30 8413 1235

Fax: +49 30 8413 1128

Email: panopoul@molgen.mpg.de

The library was characterised by oligonucleotide fingerprinting (ONFP) to reduce sequencing redundancy. According to the ONFP procedure, clones giving the same hybridisation pattern with a battery of 200 5mer oligonucleotides are grouped into clusters. One clone per cluster is selected for sequencing. The size of each cluster is an indicator of the frequency of a transcript in the analysed library. The cluster size as well the coordinates of the rest of the clones assigned to the same fingerprint cluster as the clone from which the above EST is generated is available at the amphioxus project site at: <http://www.molgen.mpg.de/amphioxus/>
 Clones and filters are distributed via the Resource Center/Primary Database of the German Genome Project (<http://www.rzpd.de>)
 PCR Primers

FORWARD: 5' CCCCAGGCTTACACTTATGCTCCGGCTCG 3' (ML13RSP)
 BACKWARD: 5' GCTATTACCCAGCTGGCGAAGGGGATG 3' (ML13SP)
 Insert Length: 1200 Std Error: 0.00
 Seq primer: 5' -CCGCTCGGAATTCGCCGT-3' pSport3/86
 High quality sequence stop: 483.

FEATURES

Location/Qualifiers
 source 1..483
 /organism="Branchiostoma floridae"
 /mol type="mRNA"
 /db xref="taxon:7739"
 /clone="MPMPG498B0638"
 /tissue type="whole embryo"
 /dev stage="5-6 hrs (gastrula stage)"
 /lab host="E.coli, XL1 blue"
 /clone lib="Amphioxus 5-6 hrs cDNA library (Name convention: BFLG or MPMPG498)"
 /note="Vector: pSport1; Site 1: SalI, KpnI, EcoRI (5'); Site 2: NotI, BamHI, HindIII (3'); OligodT primed and directionally cloned in pSport1 vector using a NotI (5'-pCACTAGTTCTAGATCGGAGCGCCGCC (T)15-3' and a SalI 5'-TCGACCCAGCGCTCCG-3' adapters (Gibco BRL)."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 483;
 Best Local Similarity 100.0%; Pred. No. 1;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 358 TAAATGAAAAA 407

RESULT 79

LOCUS CD681869

DEFINITION

Meloidogyne chitwoodi

EST.

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

1 (bases 1 to 503)

McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J., Wyite, T., Dantze, M., Marra, M., Hillier, L., Kucaba, T., Theising, B., Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C., Tsagaris, V., Ronko, I., Kennedy, S., Maguire, L., Beck, C., Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T., Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R., Waterston, R. and Wilson, R.

The Washington Univ. Nematode EST Project, 1999

Unpublished (1999)

Contact: McCarter JP

The Washington Univ. Nematode EST Project, 1999

Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA

Tel: 314 286 1800

Fax: 314 286 1810

Email: est@wustl.edu

The library was constructed by Claire Murphy and Dr. Makedonka Dautova at Washington University, St. Louis

(mdautova@wustl.edu). Oligo(dT)-SL1 PCR based library.

Meloidogyne chitwoodi egg cDNA PCR products of size 5400

nucleotides containing SL1 on the 5' end and oligo(dT) on the 3'

end were non-directionally cloned into pCRII-TOPO (Invitrogen)

following the Topo TA cloning protocol. Eggs were provided by Dr.

David Bird of North Carolina State University, Raleigh, NC

(david_bird@ncsu.edu).

Seq primer: SL1 primer

High quality sequence stop: 503.

FEATURES
source

Location/Qualifiers
1. .503
/organism="Meloioogyne chitwoodi"
/mol_type="mRNA"
/db_xref="taxon:59747"
/dev_stage="eggs"
/lab_host="DH10B"
/clone_lib="Meloioogyne chitwoodi egg SL1 TOPO v1"
/note="vector: pCR11-TOPO (Invitrogen); Site 1: EcoRI;
Site 2: EcoRI; The library was constructed by Claire
Murphy and Dr. Makedonka Bautova at Washington University,
St. Louis. Oligo(dT)-SL1 PCR based library. Meloioogyne
chitwoodi egg cDNA PCR products of size >400 nucleotides
containing SL1 on the 5' end and oligo(dT) on the 3' end
were non-directionally cloned into pCR11-TOPO(Invitrogen)
following the TOPO TA cloning protocol. Eggs were provided
by Dr. David Bird of North Carolina State University,
Raleigh, NC (david_bird@ncsu.edu)."

ORIGIN

Query Match 1.3%; Score 50; DB 14; Length 503;
Best Local Similarity 100.0%; Pred. No. 1;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAA 3762
DB 415 TAAATGAAAAA 464

RESULT 80

BM938356
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PubMed
COMMENT

BM938356 515 bp mRNA linear EST 13-MAR-2002
UI-M-BZ1-bfw-o-15-0-UI-r1 NIH BMAP MH12 S1 Mus musculus cDNA clone
UI-M-BZ1-bfw-o-15-0-UI 5', mRNA sequence.
BM938356
BM938356.1 GI:19397508
Mus musculus (house mouse)
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 515)
Ronald, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
9704477
889548
Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20892-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov
cDNA library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 83-187, >(TAA)nSimple_repeat 105-189,
>POLY(A)Simple_repeat
Seq primer: M13 REVERSE.

FEATURES
source

Location/Qualifiers
1. .515
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UI-M-BZ1-bfw-o-15-0-UI"
/dev_stage="27-32 days"
/lab_host="DH10B (Life Technologies)"

ORIGIN

/clone_lib="NIH BMAP MH12 S1"
/note="vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; The
NIH BMAP MH12 S1 library is a subtracted library derived
from NIH BMAP MH12. NIH BMAP MH12 is a library derived
from mouse hippocampus tissue. For a detailed description
of the library from which this clone was derived, please
visit our web site at brainest.eng.uowa.edu."

ORIGIN
Query Match 1.3%; Score 50; DB 12; Length 515;
Best Local Similarity 100.0%; Pred. No. 0.98;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAA 3762
DB 104 TAAATGAAAAA 153

RESULT 81

AM464287
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AM464287 562 bp mRNA linear EST 24-FEB-2000
BP230015A20F2 Soares normalized bovine placenta Bos taurus cDNA
clone BP230015A20F2 5', mRNA sequence.
AM464287
AM464287.1 GI:7034455
EST
Bos taurus (cow)
Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovidae; Bovinae; Bos.
1 (bases 1 to 562)
Lewin, H.A., Soares, M.B., Rebeiz, M., Pardinas, J., Liu, L. and
Larson, J.H.
Bovine ESTs
Unpublished (2000)
Contact: Lewin, H. A.
W. M. Keck Center for Comparative and Functional Genomics
University of Illinois at Urbana-Champaign
340 Edward R. Madigan Laboratory, 1201 W. Gregory Dr., Urbana, IL
61801, USA
Tel: 217 333 5998
Fax: 217 244 5617
Email: h-lewin@uiuc.edu

Funding for cattle EST sequencing was provided by the USDA National
Research Initiative, Animal Genome Resource Grant AG 99-3205-8534
to H. A. Lewin and J. E. Womack. Base Calling/Quality Scores: PHRED
from Washington University Genome Center. Vector trimmi g:
Cross_match from Washington University Genome Center PHRAP suite.
Sequences submitted are vector free and at least 200 bp in length.
PCR Primers
FORWARD: TAATACGACTCACTATAGG
BACKWARD: ATTAACCTCTACTAAAG
Insert Length: 562 Std Error: 0.00
Plate: BP230015A20 Row: F column: 2
Seq primer: ACGGATACAAATTCACACAGGA
High quality sequence stop: 562.

FEATURES
source

Location/Qualifiers
1. .562
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/clone="BP230015A20F2"
/sex="female"
/lab_host="DH10B"
/clone_lib="Soares normalized bovine placenta"
/note="Organ: placenta; Vector: pT73Pac; Site 1: EcoRI;
Site 2: NotI; The cDNA library was contributed by the
Soares laboratory and it was constructed and normalized
as described by Bonaldo, M.F., Lennon, G. and Soares,
M.B. (1996), Genome Research 6(9): 791-806."

```
REFERENCE
AUTHORS
  1 (bases 1 to 621)
  McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
  Wyllie, T., Dente, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
  Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
  Tsagaris, V., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
  Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
  Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
  McCann, R., Waterston, R. and Wilson, R.
  The Washington Univ. Nematode EST Project, 1999
  Unpublished (1999)
  Contact: McCarter JP
  The Washington Univ. Nematode EST Project, 1999
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@wustl.wustl.edu
  Cloned unidirectionally. Poly (A) + RNA was concentrated and
  purified using Dynabeads (Dyna) and mRNA eluted for 1st strand
  synthesis. 1st strand cDNA created using MMV RT (PowerScript,
  Clontech) and primed with oligo(dT) with XhoI site (primer,
  CDSIII/3-XhoI) and 5' SMART anchor added using chimeric DNA-RNA
  oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR
  using the SMART template was also desired. 15 PCR cycles were done
  using 1st strand and primers specific to S11 leader sequence
  (S11-NotI) and 3' end (XhoI-No-dT). dscDNA was digested using
  XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and
  ligated to digested pGEM -11zf(+) plasmid. Chemically competent
  cells were used as host. This library was provided by David Bird's
  lab (david.bird@ucsf.edu), of North Carolina State University.
  Seq primer: Sp6.
  Location/Qualifiers
  1..621
  /organism="Meloiodogyne hapla"
  /mol_type="mRNA"
  /db_xref="taxon:6305"
  /sex="female"
  /tissue_type="whole organism"
  /dev_stage="adult"
  /lab_host="DH10B"
  /clone_lib="Meloiodogyne hapla female S11 pGEM"
  /note="Vector: pGEM-11zf(+) (Promega); Site 1: XhoI;
  Site 2: NotI; Cloned unidirectionally, Poly (A) + RNA was
  concentrated and purified using Dynabeads (Dyna) and mRNA
  eluted for 1st strand synthesis. 1st strand cDNA created
  using MMV RT (PowerScript, Clontech) and primed with
  oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5'
  SMART anchor added using chimeric DNA-RNA oligo
  (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a
  PCR using the SMART template was also desired. 15 PCR
  cycles were done using 1st strand and primers specific to
  S11 leader sequence (S11-NotI) and 3' end (XhoI-No-dT).
  dscDNA was digested using XhoI/NotI, fractionated on
  Chroma-spin 400 columns (Clontech) and ligated to digested
  pGEM -11zf(+) plasmid. Chemically competent cells were
  used as host. This library was provided by David Bird's
  lab (david.bird@ucsf.edu), of North Carolina State
  University."
  Query Match 1.3%; Score 50; DB 14; Length 621;
  Best Local Similarity 100.0%; Pred. No. 0.82;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
  DB 371 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 420
  RESULT 84
  LOCUS BC049569 637 bp mRNA linear HTC 08-APR-2003
  DEFINITION Mus musculus, RIKEN cDNA 1700054O13 gene, clone IMAGE:6772969,
  Tylenchoidae; Heteroderidae; Meloiodogyninae; Meloiodogyne.
```

```
REFERENCE
AUTHORS
  1 (bases 1 to 621)
  McCarter, J., Clifton, S., Chiapelli, B., Pape, D., Martin, J.,
  Wyllie, T., Dente, M., Marra, M., Hillier, L., Kucaba, T., Theising, B.,
  Bowers, Y., Gibbons, M., Ritter, E., Bennett, J., Franklin, C.,
  Tsagaris, V., Ronko, I., Kennedy, S., Maguire, L., Beck, C.,
  Underwood, K., Steptoe, M., Allen, M., Person, B., Swaller, T.,
  Harvey, N., Schurk, R., Kohn, S., Shin, T., Jackson, Y., Cardenas, M.,
  McCann, R., Waterston, R. and Wilson, R.
  The Washington Univ. Nematode EST Project, 1999
  Unpublished (1999)
  Contact: McCarter JP
  The Washington Univ. Nematode EST Project, 1999
  Washington University School of Medicine
  4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
  Tel: 314 286 1800
  Fax: 314 286 1810
  Email: est@wustl.wustl.edu
  Cloned unidirectionally. Poly (A) + RNA was concentrated and
  purified using Dynabeads (Dyna) and mRNA eluted for 1st strand
  synthesis. 1st strand cDNA created using MMV RT (PowerScript,
  Clontech) and primed with oligo(dT) with XhoI site (primer,
  CDSIII/3-XhoI) and 5' SMART anchor added using chimeric DNA-RNA
  oligo (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a PCR
  using the SMART template was also desired. 15 PCR cycles were done
  using 1st strand and primers specific to S11 leader sequence
  (S11-NotI) and 3' end (XhoI-No-dT). dscDNA was digested using
  XhoI/NotI, fractionated on Chroma-spin 400 columns (Clontech) and
  ligated to digested pGEM -11zf(+) plasmid. Chemically competent
  cells were used as host. This library was provided by David Bird's
  lab (david.bird@ucsf.edu), of North Carolina State University.
  Seq primer: Sp6.
  Location/Qualifiers
  1..621
  /organism="Meloiodogyne hapla"
  /mol_type="mRNA"
  /db_xref="taxon:6305"
  /sex="female"
  /tissue_type="whole organism"
  /dev_stage="adult"
  /lab_host="DH10B"
  /clone_lib="Meloiodogyne hapla female S11 pGEM"
  /note="Vector: pGEM-11zf(+) (Promega); Site 1: XhoI;
  Site 2: NotI; Cloned unidirectionally, Poly (A) + RNA was
  concentrated and purified using Dynabeads (Dyna) and mRNA
  eluted for 1st strand synthesis. 1st strand cDNA created
  using MMV RT (PowerScript, Clontech) and primed with
  oligo(dT) with XhoI site (primer CDSIII/3-XhoI) and 5'
  SMART anchor added using chimeric DNA-RNA oligo
  (SMART-NotI-r-GGG). SMART-NotI-r-GGG was used in case a
  PCR using the SMART template was also desired. 15 PCR
  cycles were done using 1st strand and primers specific to
  S11 leader sequence (S11-NotI) and 3' end (XhoI-No-dT).
  dscDNA was digested using XhoI/NotI, fractionated on
  Chroma-spin 400 columns (Clontech) and ligated to digested
  pGEM -11zf(+) plasmid. Chemically competent cells were
  used as host. This library was provided by David Bird's
  lab (david.bird@ucsf.edu), of North Carolina State
  University."
  Query Match 1.3%; Score 50; DB 13; Length 602;
  Best Local Similarity 100.0%; Pred. No. 0.84;
  Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
  QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
  DB 386 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 435
  RESULT 83
  LOCUS CF979621 621 bp mRNA linear EST 25-NOV-2003
  DEFINITION r979a01.y1 Meloiodogyne hapla female S11 pGEM Meloiodogyne hapla cDNA
  5', mRNA sequence.
  ACCESSION CF979621
  VERSION CF979621.1 GI:38512670
  KEYWORDS EST.
  SOURCE Meloiodogyne hapla
  ORGANISM Meloiodogyne hapla
  Eukaryota; Metazoa; Nematoda; Chromadorea; Tylenchida; Tylenchina;
  Tylenchoidae; Heteroderidae; Meloiodogyninae; Meloiodogyne.
```

```

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM14822 row: d column: 06
High quality sequence stop: 103.
Location/Qualifiers
1..780
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7050776"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC 7). Library was constructed by Open Biosystems
(Huntsville, AL)."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 780;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 61 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 86
LOCUS BUS98654 844 bp mRNA linear EST 20-SEP-2002
DEFINITION AGENCOURT_8822419 NIH_MGC_142 Homo sapiens cDNA clone IMAGE:6457561
5', mRNA sequence.
ACCESSION BUS98654
VERSION BUS98654.1 GI:23250413
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2639 row: k column: 02

mRNA
BC049569
VERSION BC049569.1 GI:29612485
KEYWORDS HTC.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 637)
REFERENCE
AUTHORS Strausberg, R.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK
COMMENT NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-remail.nih.gov
Tissue Procurement: Dr. Jonathan Kuo, NIMH
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steven Jones, Jennifer Asano, Ian Bosdet, Yaron Butterfield,
Susanna Chan, Readman Chiu, Chris Fjell, Erin Garland, Ran Guin,
Leticia Hsiao, Martin Krzywinski, Reta Kutsche, Oliver Lee, Soo
Sen Lee, Victor Ling, Carrie Mathewson, Candice McLeavy, Steven
Nees, Pawan Pandoh, Anna-Liisa Prabhu, Parvaneh Saedi, Jacqueline
Schein, Duane Smalius, Michael Smith, Lorraine Spence, Jeff Stott,
Michael Thorne, Miranada Tsai, Natasja van den Bosch, Jill Vardy,
George Yang, Scott Zuyderduyn, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 46 Row: i Column: 12
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 13385599
This clone has the following problem: retained intron.
Location/Qualifiers
1..637
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6772969"
/tissue_type="Testicle, mouse"
/clone_lib="NIH_MGC_169"
/lab_host="DH10B"
/notes="Vector: pDNR-LIB"

ORIGIN
Query Match 1.3%; Score 50; DB 11; Length 637;
Best Local Similarity 100.0%; Pred. No. 0.8;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 507 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 556

RESULT 85
LOCUS CK025580 780 bp mRNA linear EST 26-NOV-2003
DEFINITION AGENCOURT_16619085 NIH_ZGC_10 Danio rerio cDNA clone IMAGE:7050776
5', mRNA sequence.
ACCESSION CK025580
VERSION CK025580.1 GI:38551504
KEYWORDS EST.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.

```

```

REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Daniela S. Gerhard, Ph.D.
Office of Cancer Genomics
National Cancer Institute / NIH
Bldg. 31 Rm10A07 Bethesda, MD 20892
Email: cgapbs-remail.nih.gov
Tissue Procurement: Len Zon, Harvard
cDNA Library Preparation: Open Biosystems
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM14822 row: d column: 06
High quality sequence stop: 103.
Location/Qualifiers
1..780
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="IMAGE:7050776"
/tissue_type="whole body"
/lab_host="DH10B"
/clone_lib="NIH_ZGC_10"
/notes="Vector: pExpress1; Site 1: NotI; Site 2: EcoRV;
Bulk tissue was collected from a whole adult individual
from the Tuebingen strain. 1st strand cDNA was primed with
a Not I - oligo(dT) primer, double-stranded cDNA was
cloned into the Not I and EcoRV sites of pExpress-1.
Library was size-selected for >1 kb fragments. A
normalized version of this library is also available
(NIH_ZGC 7). Library was constructed by Open Biosystems
(Huntsville, AL)."

ORIGIN
Query Match 1.3%; Score 50; DB 14; Length 780;
Best Local Similarity 100.0%; Pred. No. 0.65;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
Db 61 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 12

RESULT 86
LOCUS BUS98654 844 bp mRNA linear EST 20-SEP-2002
DEFINITION AGENCOURT_8822419 NIH_MGC_142 Homo sapiens cDNA clone IMAGE:6457561
5', mRNA sequence.
ACCESSION BUS98654
VERSION BUS98654.1 GI:23250413
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
AUTHORS NIH-MGC http://mgc.nci.nih.gov/
TITLE NIH-MGC Consortium of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW2639 row: k column: 02

```

High quality sequence stop: 403.

FEATURES

source

1. .844
Location/Qualifiers
/organism="Homo sapiens"
/db_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:6457561"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (TI-phage-resistant)"
/clone_lib="NIH_MGC_142"
/note="Vector: pDNR-LIB; Site 1: SfiI (ggccattatggcc); Site 2: SfiI (ggcgcctctggcc); Double-stranded cDNA was prepared from a pool of 40 cell line polyA+ RNAs (bladder - 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon - 4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%, kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%, ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were used in cloning as follows:
5'-AGCAGTGCTATCACGACGAGTGGCATTCGCGCGG-3' and 5'-ATTCTAGGCGCGGCGGCGGACATG-dT(30)NN-3'. Full-length enriched library was constructed using the Clontech Creator SMART kit and size-selected to contain the >0.5 kb size fraction (other fractions present in NIH_MGC_141). Library created in the laboratory of M. Brownstein (NIH). Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.3%; Score 50; DB 13; Length 844;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 849 bp mRNA linear EST 22-NOV-2000

Db 363 TAAATGAAAAA 849 bp mRNA linear EST 22-NOV-2000

RESULT 87
BF339838
LOCUS
DEFINITION
602038845F1 NCI_CGAP_Brn64 Homo sapiens cDNA clone IMAGE:4186834
5', mRNA sequence.

ACCESSION
BF339838

VERSION
BF339838.1 GI:11286297

KEYWORDS
EST.

SOURCE
Homo sapiens (human)

ORGANISM
Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 849)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: David N. Louis, M.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLN9508 row: i column: 11

High quality sequence stop: 568.

Location/Qualifiers

1. .849

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:4186834"

/tissue_type="globlastoma with EGFR amplification"

/lab_host="DH10B (TI phage-resistant)"

/clone_lib="NCI_CGAP_Brn64"

ORIGIN

Query Match 1.3%; Score 50; DB 10; Length 849;
Best Local Similarity 100.0%; Pred. No. 0.6;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 865 bp mRNA linear EST 29-SEP-2000

Db 485 TAAATGAAAAA 865 bp mRNA linear EST 29-SEP-2000

RESULT 88

BE913663

LOCUS

DEFINITION

601669014F1 NCI_CGAP_Mam1 Mus musculus cDNA clone IMAGE:3969146 5',

mRNA sequence.

ACCESSION

BE913663

VERSION

BE913663.1 GI:10411508

KEYWORDS

EST.

SOURCE

Mus musculus (house mouse)

ORGANISM

Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 865)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Robert Strausberg, Ph.D.

Email: cgapbs-remail.nih.gov

Tissue Procurement: Gilbert Smith, Ph.D.

cDNA Library Preparation: Life Technologies, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: LLN9146 row: k column: 03

High quality sequence stop: 406.

Location/Qualifiers

1. .865

/organism="Mus musculus"

/mol_type="mRNA"

/strain="FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:3969146"

/tissue_type="tumor, biopsy sample"

/dev_stage="10 months, virgin"

/lab_host="DH10B"

/clone_lib="NCI_CGAP Mam1"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site 1: SalI;

Site 2: NotI; Cloned unidirectionally. Primer: Oligo dT.

Library constructed by Life Technologies. Investigator

providing samples: Gilbert Smith, NIH"

Query Match 1.3%; Score 50; DB 10; Length 865;

Best Local Similarity 100.0%; Pred. No. 0.59;

Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 883 bp mRNA linear EST 05-MAY-2003

Db 353 TAAATGAAAAA 883 bp mRNA linear EST 05-MAY-2003

RESULT 89

EX346858

LOCUS

DEFINITION

EX346858 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA

clone CSODI005YF20 5-PRIME, mRNA sequence.

```

ACCESSION BX346858
VERSION BX346858.1 GI:30379137
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
1 (bases 1 to 883) Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
JOURNAL Full-length cDNA libraries and normalization
COMMENT Unpublished (2001)
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AT0172E12_T01626_1.
FEATURES
    source
    1..883
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CS0D1005YF20"
    /tissue_type="PLACENTA COT 25-NORMALIZED"
    /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
    /notes="1st strand cDNA was primed with a NotI-oligo(dT)
    primer. Five prime end enriched, double-strand cDNA was
    digested with Not I and cloned into the Not I and EcoR V
    sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 883;
Best Local Similarity 100.0%; Pred. No. 0.58;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 916 bp mRNA linear EST 12-JUL-2000
||||| 3762
Db 28 TAAATGAAAAA 77

RESULT 90
LOCUS AU051237 892 bp mRNA linear EST 12-JUL-2000
DEFINITION AU051237 Sugano mouse brain mmb Mus musculus cDNA clone MNCB-1712
5', mRNA sequence.
ACCESSION AU051237
VERSION AU051237.1 GI:4434246
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
23-1, Toyama I-chome, Shinjuku-ku, Tokyo 162-8640, Japan
Hata, H., Yanaguchi, R., Watanabe, M., Imai, J., Shibui, A., Yoshida, K.,
Sasaki, M., Suzuki, Y., Tatematsu, S. and Sugano, S.
TITLE Construction of mouse full length-enriched cDNA libraries by
oligo-capping
JOURNAL Unpublished (1999)
COMMENT Contact: Katsuyuki Hashimoto
Division of Genetic Resources
National Institute of Infectious Diseases
23-1, Toyama I-chome, Shinjuku-ku, Tokyo 162-8640, Japan
Email: khashi@nih.go.jp
URL: http://www.nih.go.jp/yoken/genbank/
Seq primer: 5' end primer: CTTCTGCTCTAAAGCTGGG
POLYA=No.
FEATURES
    source
    1..892
    /organism="Mus musculus"
    /mol_type="mRNA"
    /strain="C57BL"

```

```

/db_xref="taxon:10090"
/clone="MNCB-1712"
/sex="female"
/dev_stage="adult"
/lab_hosts="TOP10"
/clone_lib="Sugano mouse brain mmb"
/note="Organ: brain; Vector: pME18S-FL3; 1st strand cDNA
was primed with an oligo(dT) primer
ATGTGGCTTTTCTTTTCTTTT; double-stranded cDNA was
ligated to a DraIII adaptor [TGTTGGCTACTGG], digested and
cloned into distinct DraIII sites of the pME18S-FL3. XhoI
sites just outside the DraIII sites can be used to isolate
the cDNA insert. Size selection was performed to exclude
fragments <1.5 kb. Library was constructed by Sugano et
al. (University of Tokyo, Institute of Medical Science).
Custom primer used for sequencing: 5' end primer
[CTTCTGCTCTAAAGCTGGG], 3' end primer
[CGACCTGCAGCTCGAGCACA]"
ORIGIN
Query Match 1.3%; Score 50; DB 9; Length 892;
Best Local Similarity 100.0%; Pred. No. 0.57;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA 916 bp mRNA linear EST 01-MAY-2003
||||| 3762
Db 116 TAAATGAAAAA 165

RESULT 91
LOCUS BX325997 916 bp mRNA linear EST 01-MAY-2003
DEFINITION BX325997 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CS0D1082YI06 5-PRIME, mRNA sequence.
ACCESSION BX325997
VERSION BX325997.1 GI:30309528
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 916)
Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7394.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AT024ZC12_T02299_1&cluster=7394.f.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AT024ZC12_T02299_1.
FEATURES
    Location/Qualifiers
    1..916
    /organism="Homo sapiens"
    /mol_type="mRNA"
    /db_xref="taxon:9606"
    /clone="CS0D1082YI06"
    /tissue_type="PLACENTA COT 25-NORMALIZED"
    /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
    /note="1st strand cDNA was primed with a NotI-oligo(dT)
    primer. Five prime end enriched, double-strand cDNA was
    digested with Not I and cloned into the Not I and EcoR V
    sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 916;
Best Local Similarity 100.0%; Pred. No. 0.56;

```

```
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 44 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 93

RESULT 92
BX369283
LOCUS
DEFINITION
BX369283 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI054YML7 5-PRIME, mRNA sequence.
ACCESSION
BX369283
VERSION
BX369283.1 GI:30441388
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 940)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 4928.r For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAF027ZH10_AF02577.1&cluster=4928.r.
Contact : Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0BAF027ZH10_AF02577.1.
Location/Qualifiers
1..940
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CSODI054YML7"
/tissue_type="PLACENTA COT 25-NORMALIZED"
/clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
/Note="5' end strand cDNA was primed with a NotI-oligo (dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 940;
Best Local Similarity 100.0%; Pred.No.0.54;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 209 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 258

RESULT 93
BU945029
LOCUS
DEFINITION
AGENCYCOURT 10546264 NIH_MGC_141 Homo sapiens cDNA clone
IMAGE:673277 5', mRNA sequence.
ACCESSION
BU945029
VERSION
BU945029.1 GI:24133848
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 942)
NIH-MGC http://mgc.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
```

```
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
Tissue Procurement: NCI
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLCW3062 row: k column: 04
High quality sequence stop: 352.
Location/Qualifiers
1..942
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:673277"
/tissue_type="mixed (pool of 40 RNAs)"
/lab_host="DH10B (T1-phage-resistant)"
/clone_lib="NIH_MGC_141"
/Note="Vector: pDNR-LIB; Site: 1: SfiI (ggccattatggcc);
Site 2: SfiI (ggcgctcgcc); Double-stranded cDNA was
prepared from a pool of 40 cell line polyA+ RNAs (bladder
- 2%, blood - 33.4%, brain - 5.6%, breast - 12.5%, colon -
4%, connective tissue - 1.4%, eye - 1%, intestine - 2.6%,
kidney - 2.2%, liver - 5.7%, lung - 10.8%, NK-cell - 5.2%,
ovary - 4%, pharynx - 2.5%, prostate - 4.3%, salivary
gland - 1.3%, and skin - 2.3%). 5' and 3' adaptors were
used in cloning as follows:
5'-AAGCAGTGTATCAACGACGAGTGGCCATTACGCGGG-3' and
5'-ATTCTAGAGCGGCGGCGGCGCATG-dt(30)NN-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.2-0.5
kb size fraction (other fractions present in NIH_MGC 142).
Library created in the laboratory of M. Brownstein (NIH,
NIH). Note: this is a NIH_MGC library."

ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 942;
Best Local Similarity 100.0%; Pred.No.0.54;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3713 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
    |||||
Db 245 TAAATGAAAAA...AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 294

RESULT 94
BX326372
LOCUS
DEFINITION
BX326372 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
clone CSODI082YI06 5-PRIME, mRNA sequence.
ACCESSION
BX326372
VERSION
BX326372.1 GI:30344535
KEYWORDS
SOURCE
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 942)
Li, W.B., Gruber, C., Jessee, J., and Polayes, D.
Full-length cDNA libraries and normalization
Unpublished (2001)
JOURNAL
COMMENT
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 7394.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0BAA018ZE02_CS01646_1&cluster=7394.f.
```

Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAA0182E02_CS01646_1.
 Location/Qualifiers

FEATURES
 source
 1..942
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1082Y106"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 1.3%; Score 50; DB 13; Length 942;
 Best Local Similarity 100.0%; Pred. No. 0.54; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0

Qy 3713 TAAATGAAAAA... 952 bp mRNA linear EST 08-MAY-2003
 Db 122 TAAATGAAAAA... 3762

RESULT 95
 LOCUS
 DEFINITION
 BX369284 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1054YM17 5-PRIME, mRNA sequence.
 ACCESSION
 BX369284
 VERSION
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 952)
 Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 4928.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAF027ZH10_AF02577_2&cluster=4928.r.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAF027ZH10_AF02577_2.
 Location/Qualifiers

FEATURES
 source
 1..952
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1054YM17"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 1.3%; Score 50; DB 13; Length 952;
 Best Local Similarity 100.0%; Pred. No. 0.53; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0

Qy 3713 TAAATGAAAAA... 952 bp mRNA linear EST 05-FEB-2002
 Db 123 TAAATGAAAAA... 172

Db 211 TAAATGAAAAA... 260

RESULT 96
 LOCUS
 DEFINITION
 BX369858 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
 clone CS0D1006YN18 5-PRIME, mRNA sequence.
 ACCESSION
 BX369858
 VERSION
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 956)
 Li.W.B., Gruber.C., Jessee.J. and Polayes.D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 COMMENT
 Contact: Genoscope
 Genoscope - Centre National de Sequencage
 BP 191 91006 EVRY cedex - France
 Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 Library was constructed by Life Technologies, a division of
 Invitrogen. This sequence belongs to sequence cluster 3779.r For
 more information about this cluster, see
 http://www.genoscope.cns.fr/
 cgi-bin/cluster.cgi?seq=CS0BAG025ZA03_CS02321_1&cluster=3779.r.
 Contact : Feng Liang Email : fliang@lifetech.com URL :
 http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
 Faraday Avenue Genoscope sequence ID : CS0BAG025ZA03_CS02321_1.
 Location/Qualifiers

FEATURES
 source
 1..956
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0D1006YN18"
 /tissue_type="PLACENTA COT 25-NORMALIZED"
 /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
 /note="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and cloned into the Not I and EcoR V
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Query Match 1.3%; Score 50; DB 13; Length 956;
 Best Local Similarity 100.0%; Pred. No. 0.53; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0

Qy 3713 TAAATGAAAAA... 3762
 Db 123 TAAATGAAAAA... 172

RESULT 97
 LOCUS
 DEFINITION
 BM459621 AGENCOURT 6417960 NIH_MGC_71 Homo sapiens cDNA clone IMAGE:5534247
 5', mRNA sequence.
 ACCESSION
 BM459621
 VERSION
 EST.
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 994)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-r@mail.nih.gov

Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM12220 row: 9 column: 16
 High quality sequence start: 28
 High quality sequence stop: 412.
 Average insert size 2.1 kb.

FEATURES

1. 994
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:553424"
 /tissue_type="leiomyosarcoma"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 71"
 /note="Organ: uterus; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
 Average insert size 2.1 kb."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 994;
 Best Local Similarity 100.0%; Pred. No. 0.51;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 255 TAAATGAAAAA 304

RESULT 98

BG254352
 LOCUS
 DEFINITION
 602368921F1 NTH_MGC_91 Homo sapiens cDNA clone IMAGE:4477133 5',
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: DCTD/DTF
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10305 row: 1 column: 06
 High quality sequence stop: 196.

FEATURES

1. 1023
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4477133"
 /tissue_type="adenocarcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 91"
 /note="Organ: Prostate; Vector: pCMV-SPORT6; Site 1: NotI;
 Site 2: SalI; Cloned unidirectionally; oligo-dt primed.
 Average insert size 1.4 kb. Library enriched for
 full-length clones and constructed by Life Technologies.

Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 1023;
 Best Local Similarity 100.0%; Pred. No. 0.5; 0; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 139 TAAATGAAAAA 188

RESULT 99

BG105078
 LOCUS
 DEFINITION
 602312366F1 NTH_MGC_84 Homo sapiens cDNA clone IMAGE:4422088 5',
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 COMMENT
 NIH-MGC <http://mgc.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM10162 row: c column: 17
 High quality sequence stop: 444.

FEATURES

SOURCE

1. 1041
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4422088"
 /tissue_type="adrenal cortex carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH MGC 84"
 /note="Organ: adrenal gland; Vector: pCMV-SPORT6; Site 1:
 NotI; Site 2: SalI; Cloned unidirectionally; oligo-dt
 primed. Average insert size 1.229 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Query Match 1.3%; Score 50; DB 12; Length 1041;
 Best Local Similarity 100.0%; Pred. No. 0.49; 0; Indels 0; Gaps 0;
 Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3713 TAAATGAAAAA 3762
 |||||
 Db 351 TAAATGAAAAA 400

RESULT 100

BUS26227
 LOCUS
 DEFINITION
 AGENCOURT 10173174 NIH MGC 144 Mus musculus cDNA clone
 IMAGE:6534925 5', mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

Mus musculus (house mouse)

```

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.
REFERENCE 1 (bases 1 to 1070)
AUTHORS NIH-MGC http://mgs.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. Michael Brownstein
cDNA Library Preparation: Michael Brownstein Laboratory
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1CM2693 row: n column: 13
High quality sequence stop: 422.
FEATURES
Location/Qualifiers
1..1070
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/clone="IMAGE:6534925"
/lab_host="DH10B (Ti-phage-resistant)"
/clone_lib="NIH_MGC_144"
/note="Organ: Brain; Vector: pDNR-LIB; Site 1: SfiI
(ggcatataggcc); Site 2: SfiI (ggcgctgggc); cDNA made
by oligo-dT priming and directionally cloned. 5' and 3'
adaptors were used in cloning as follows:
5'-AAGCAGTGTATCAACGACGATGCGCATTCAGCGCGG-3' and
5'-ATTCTAGAGCGGCGGCGGCGGCGG-3'. Full-length
enriched library was constructed using the Clontech
Creator SMART kit and size-selected to contain the 0.2-0.5
kb size fraction (other fractions present in NIH_MGC_143).
Library created in the laboratory of M. Brownstein (NIMH,
NIH). Note: this is a NIH_MGC Library."
ORIGIN
Query Match 1.3%; Score 50; DB 13; Length 1070;
Best Local Similarity 100.0%; Pred. No. 0.48;
Matches 50; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3713 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 3762
| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 525 TAAATGAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 574

Search completed: April 23, 2004, 11:01:47
Job time : 8962 secs

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 22, 2004, 16:09:21 ; Search time 2165 Seconds
(without alignments)
2702.683 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715

Sequence: 1 MGNVLRKLSCLGENQKPK.....SVSRPCSTHEHDYEVVPH 135

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DSV=xlp
-Q=/cgn2 1/USPTO.spool_p/US10005907/runat_21042004_113754_11432/app.query.fasta_1.327
-DB=GenEmbl -QFMT=fastap -SUFFIX=se -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNIT=bits -START=1 -END=1 -NATRIK=blosum62 -TRANS=human40.cgi -LIST=45
-OUTFMT=ptc -THR SCORE=ptc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-USER=US10005907@cgn 1 1 5265 @runat_21042004_113754_11432 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.in.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*
15: em.ba.*
16: em.fun.*
17: em.hum.*
18: em.in.*
19: em.mu.*
20: em.om.*
21: em.or.*
22: em.ov.*
23: em.pat.*
24: em.ph.*
25: em.pl.*
26: em.io.*
27: em.sts.*
28: em.un.*

29: em.vi.*
30: em.htg.hum.*
31: em.htg.inv.*
32: em.htg.other.*
33: em.htg.mus.*
34: em.htg.pln.*
35: em.htg.rod.*
36: em.htg.mam.*
37: em.htg.vrt.*
38: em.sy.*
39: em.htg.hum.*
40: em.htg.mus.*
41: em.htg.other.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	2648	9	BC024174 Homo sapi
2	715	100.0	3762	6	AX505122 Sequence
3	711	99.4	1709	9	AX123798 Homo sapi
4	634	88.7	1924	9	AB060908 Macaca fa
5	584	81.7	4032	9	AX126682 Homo sapi
6	581	81.3	4136	9	AX124520 Homo sapi
c	430	60.1	141268	2	AC074365 Homo sapi
8	430	60.1	185467	9	AL606804 Human DNA
9	169.5	23.7	1532	9	BC030506 Homo sapi
10	169.5	23.7	1654	9	AF521911 Homo sapi
11	169.5	23.7	3270	9	AF521911 Homo sapi
12	159	22.2	1643	6	AX740457 Sequence
13	128	17.9	1672	10	U13263 Mus musculu
14	126	17.6	4218	9	AX125521 Homo sapi
15	126	17.6	81001	9	AC128688 Homo sapi
16	126	17.6	203405	2	AC024964 Homo sapi
17	126	17.6	214324	2	AC110870 Homo sapi
18	112	15.7	86826	3	PFMA13P5 Plasmodu
19	110.5	15.5	212287	2	AC108574 Rattus no
20	109.5	15.3	184705	2	AC104910 Mus muscu
21	109.5	15.3	202125	2	AC124636 Mus muscu
22	104.5	14.6	78001	2	AC139696 Homo sapi
23	102.5	14.3	18920	10	AL954353 Mouse DNA
24	102	14.3	278943	2	AC147571 Pan trogl
25	100	14.0	190641	10	AL626775 Mouse DNA
26	100	14.0	286358	2	AL713872 Mus muscu
27	99.5	13.9	139117	2	AC068147 Homo sapi
28	98.5	13.8	34984	3	AC005801 Leishmani
29	98.5	13.8	35573	3	AC009603 Leishmani
30	98.5	13.8	108797	3	AC121231 Leishmani
31	98	13.7	1892	6	BD136402 95 human
32	97.5	13.6	162153	9	AC092331 Homo sapi
33	97.5	13.6	169794	2	AC004688 Plasmodu
34	97.5	13.6	188887	9	AC116364 Homo sapi
35	97.5	13.6	189648	2	AC069163 Homo sapi
36	97.5	13.6	196149	2	AC004709 Plasmodu
37	97.5	13.6	252650	3	AE014847 Plasmodu
38	97	13.6	238748	2	AC120746 Rattus no
39	97	13.6	244341	2	AC108226 Rattus no
40	96.5	13.5	156698	10	AL837509 Mouse DNA
41	96.5	13.5	175340	2	AC110159 Mus muscu
42	96.5	13.5	175754	2	AC110189 Mus muscu
43	96.5	13.5	204081	2	AC130681 Mus muscu
44	96.5	13.5	231930	2	AC134911 Mus muscu
45	96.5	13.5	249025	2	AC131475 Rattus no

ALIGNMENTS

RESULT 1

```

BC024174      2648 bp      mRNA      linear      PRI 410-OCT-2003
LOCUS        Homo sapiens LOC148823, mRNA (CDNA clone MGC:24564 IMAGE:4109064),
complete cds.
BC024174      2648 bp      mRNA      linear      PRI 410-OCT-2003
LOCUS        Homo sapiens LOC148823, mRNA (CDNA clone MGC:24564 IMAGE:4109064),
complete cds.
ACCESSION    BC024174
VERSION      BC024174.1 GI:18848218
KEYWORDS     MGC.
SOURCE       Homo sapiens (human)
ORGANISM     Homo sapiens
REFERENCE    1 (bases 1 to 2648)
AUTHORS      Krausberg,R.L., Feingold,E.A., Grouse,L.H., Derge,J.G.,
Klausner,R.D., Collins,F.S., Wagner,F.H., Shennan,C.M., Schuler,G.D.,
Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K.,
Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F.,
Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L.,
Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L.,
Scheetz,T.E., Brownstein,M.J., Usdin,T.B., Toshiyuki,S.,
Carninci,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J.,
Abramson,R.D., Mullany,S.G., Bosak,S.A., McEwan,P.J.,
McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S.,
Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W.,
Villalón,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A.,
Fahey,J., Helton,E., Kettman,M., Madan,A., Young,A.C., Shevchenko,Y.,
Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y.,
Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D.,
Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M.,
Butterfield,S., Krzywinski,M.I., Skalski,U., Smalios,D.E.,
Schnerch,A., Schein,J.B., Jones,S.J. and Marra,M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388257
12477932
JOURNAL      Strausberg,R.
MEDLINE      Direct Submission
PUBMED       Submitted (19-FEB-2002) National Institutes of Health, Mammalian
AUTHORS      Gene Collection (MGC), Cancer Genomics Office, National Cancer
TITLE         Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
JOURNAL      USA
REMARK       NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT       Contact: MGC help desk
              Email: cgapbs-remail.nih.gov
              Tissue Procurement: ATCC
              CDNA Library Preparation: CLONTECH Laboratories, Inc.
              CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILLNI)
              DNA Sequencing by: Institute for Systems Biology
              http://www.systemsbio.org
              contact: amadaneysystemsbiology.org
              Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha
              Madan, Stephanie Rodrigues, Amy Sanchez and Michelle Whiting
              Clone distribution: MGC clone distribution information can be found
              through the I.M.A.G.E. Consortium/ILLNI at: http://image.llnl.gov
              Series: IRAL Places: 32 Row: b Column: 10
              This clone was selected for full length sequencing because it
              passed the following selection criteria: matched mRNA gi: 21687148.
              Location/Qualifiers
                1..2648
                  /organism="Homo sapiens"
                  /mol_type="mRNA"
                  /db_xref="taxon:9606"
                  /clone="MGC:24564 IMAGE:4109064"
                  /tissue_type="Bone marrow, acute myelogenous leukemia"
                  /clone_lib="NIH MGC_55"
                  /lab_host="DH10B"
                  /note="Vector: pDNR-LIB"
                1..2648
                  /gene="LOC148823"
                  /db_xref="LocusID:148823"
                48..455
                  /db_xref="LocusID:148823"
                  /note="unassigned DNA"
                  /db_xref="taxon:9606"
                  /note="unassigned protein product"
                  /codon_start=1
                  /product="LOC148823"
                  /protein_id="AAH24174.1"
                  /db_xref="GI:18848219"
                  /db_xref="LocusID:148823"
                  /translation="MGNYLLKSLCLGENOKPKKGNPDRERKQEMTTPFKLQDOD
                  KKSQEVASTSQNENSGSEEVCTVYVNIHPQRSLSLSDNDQYENDISLTRVRF
                  RERSETAYLLKTSVSRPCSTHEHDYVFPF"
              Alignment Scores:
                Pred. No.: 2,12e-65      Length: 2648
                Score: 715.00      Matches: 135
                Percent Similarity: 100.00%      Conservative: 0
                Best Local Similarity: 100.00%      Mismatches: 0
                Query Match: 100.00%      Indels: 0
                DB: 9      Gaps: 0
              US-10-005-907-2 (1-135) x BC024174 (1-2648)
              QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysValprolys-20
              Db 48 ATGGGAATATTCCTCTCGAAATCTAGTTGCTTGGGAGAGAAATCAAAAGAACCCCAAG 107
              QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
              Db 108 AAAGGAACCCAGATGAGGAAAGAAACGGCAGGAATGACTACATTTGAAAGAAACTT 167
              QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
              Db 168 CAAGATCAAGATAGAAAGCCAGAGATTTTCATCCATCTTCAATCAGGAAACGAGAAAT 227
              QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
              Db 228 GCGATGGTTCGAGAGAGTGTCTACTCTCATTCATCATCCCCCATCAGATCC 287
              QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
              Db 288 TCCTCGAGTCCCAATCATGATGCTGCTATGAGAACATTCCTCCCTCACAAGGAAAGTGAGA 347
              QY 101 GlnPheArgGluArgSerGluThrGluTyrValAlaLeuLeuArgThrSerValSerArgPro 120
              Db 348 CAGTTAGAGAAAGTTCAGACAGACAGATATGCCCTCTTAGGACTTCTGTAGAGGCT 407
              QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
              Db 408 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC 452
              RESULT 2
              AX505122      3762 bp      DNA      linear      PAT 27-SEP-2002
              LOCUS        Sequence 1 from Patent WO0246389.
              DEFINITION    AX505122
              ACCESSION    AX505122
              VERSION     AX505122.1 GI:23386429
              KEYWORDS     Homo sapiens (human)
              SOURCE      Homo sapiens
              ORGANISM     Homo sapiens
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
              Nucleotide sequence with allergic hypersensitivity and mast cell
              activation
              Patent: WO 0246389-A 1 13-JUN-2002;
              DUB, S.A. (BE)
              Location/Qualifiers
                1..3762
                  /organism="Homo sapiens"
                  /mol_type="unassigned DNA"
                  /db_xref="taxon:9606"
                25..432
                  /note="unassigned protein product"
                  /codon_start=1
              CDS

```

/protein_id="CAD48808.1"
/db_xref="GI:23386430"
/db_xref="RENTREML:CAD48808"
/translation="MGVILLRLSLGLENQKPKKNGDEERKEQMTTTPERKLQDOD
KKSQEVSSNENSGSSEVCYVINHPHORSLSLSDGVDENIDSLTRKVRQF
RESETEYALLATSVSRPCSCTHEHDYEVVFPH"

ORIGIN

Alignment Scores:

Pred. No.: 3,126-65 Length: 3762
Score: 715.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x AX505122 (1-3762)

QY 1 MetGlyAsnTyrIleuLeuArgLysLeuSerCysLeuGluAsnGlnLysProLys 20
DB 25 ATGGGAATATCTCTCGGAAACTCAGTTCCTGGGAGAGATCAAAAGAGCCCAAG 84
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 85 AAGGAAACCCAGATGAGGAAAGAAACCGGAGGAATGACTACATTTGAAGAAACTT 144
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
DB 145 CAAGATCAAGATAAGAAACCCAGGAAGTTTCATCCACTTCTAATCAGGAAACGAGAT 204
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 205 GGCAGTGGTCTGGAAGTGGTGTACACTGTCATTAATCATCCCTCCATCAGAGATCC 264
QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 265 TCCTCAGCTCCCAATGATGCTATGAGAACATTCCTCCTCACAAGGAAAGTGAGA 324
QY 101 GlnPheArgLysSerGluThrGluTyrAlaLeuArgThrSerValSerArgPro 120
DB 325 CAGTTTAGAAGAGGTCAGAGACAGAAATGCGCTCTTAGGACTCTCTGTAGAGGCT 384
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 385 TGTTCTGCACCATGAGCATGATTATGAAGTTGTGTTCACAC 429

RESULT 3

AK123798 1709 bp mRNA linear PRI 09-SEP-2003
LOCUS Homo sapiens cDNA FLJ41804 fis, clone NOV2000710.
DEFINITION AK123798
VERSION AK123798.1 GI:34529425
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

1 Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,
Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R.,
Yoshikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H.,
Onogawa, S., Kaeriyama, S., Satoh, N., Matsumura, H., Takahashi, E.,
Kataoka, R., Kuga, N., Kuroda, A., Satoh, I., Kamata, K., Takami, S.,
Teraehina, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T.,
Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Wagatsuna, M., Murakawa, Y., Sekine, M., Kikuchi, K., Kanda, K.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagabari, K., Masuno, Y., Nagai, K. and Isogai, T.

NEDO human cDNA sequencing project

TITLE

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1709)

Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing;
Research Association for Biotechnology (RAB); cDNA library
construction; Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- and 3'-end one pass sequencing; RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing; HRI and
RAB; annotation: HRI and RAB.

COMMENT

FEATURES

source

1..1709
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NOVAR2000710"
/tissue_type="ovary"
/clone_lib="NOVAR2"
/dev_stage="adult"
/note="Cloning vector: pME18SFL3"

ORIGIN

Alignment Scores:

Pred. No.: 3,466-65 Length: 1709
Score: 711.00 Matches: 134
Percent Similarity: 100.00% Conservative: 1
Best Local Similarity: 99.26% Mismatches: 0
Query Match: 99.44% Indels: 0
DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AK123798 (1-1709)

QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGluAsnGlnLysProLys 20
DB 59 ATGGGAATATCTCTCGGAAACTCAGTTCCTGGGAGAGATCAAAAGAGCCCAAG 118
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 119 AAGGAAACCCAGATGAGGAAAGAAACCGGAGGAATGACTACATTTGAAGAAACTT 178
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
DB 179 CAAGATCGAGATAAGAAACCCAGAGATTTTCATCCACTTCTAATCAGGAAACGAGAT 238
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 239 GGCAGTGGTCTGGAAGAGTGTGTACACTGTCTAATTAATCATCCCTCCATCAGAGATCC 298
QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 299 TCCTCAGCTCCCAATGATGCTATGAGAACATTCCTCCTCACAAGGAAAGTGAGA 358
QY 101 GlnPheArgLysSerGluThrGluTyrAlaLeuArgThrSerValSerArgPro 120
DB 359 CAGTTTAGAAGAGGTCAGAGACAGAAATGCGCTCTTAGGACTCTCTGTAGAGGCT 418
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 419 TGTTCTGCACCATGAGCATGATTATGAAGTTGTGTTCACAC 463

RESULT 4

LOCUS

DEFINITION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

Unpublished

REFERENCE

2 (bases 1 to 1709)


```

Best Local Similarity: 85.19% Mismatches: 0
Query Match: 81.68% Indels: 20
DB: 9 Gaps: 1

US-10-005-907-2 (1-135) x AK126682 (1-4032)

QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 377 ATGGGAATATCTCTCGCAAACTC----- 403
QY 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 404 -----AGCAGGAATGACTACATTGGAAGAACTT 436
QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsn 60
DB 437 CAAGATCAAGATAAGAAAGCAAGAGTTTCACACTTCTAATCAGGAAACGAGAAT 496
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 497 GGCAAGTGGTCTGGAAGAGTGCTGACACTGTCATTATCATATCCCATCCAGATCC 556
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 557 TCCTGAGTCCCATGATGATGGCTATGAGACATTCACCTCCCAAGAAAGTGAGA 616
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 617 CAGTTTAGAAGAAAGGTGAGACAGAGATATGCCCTTCTTAGGACTTCTGTAGAGCCT 676
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 677 TGTTCCTGCACCATGAGCATGATTATGAAGTTGTGTTCACAC 721

RESULT 6
AK124520 4136 bp mRNA linear PRI 09-SEP-2003
LOCUS Homo sapiens cDNA FLJ42529 fis, clone BRACE3002390.
ACCESSION AK124520
VERSION AK124520.1 GI:34530324
KEYWORDS oligo capping, fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Ninomiya K., Wagatsuma M., Kanda K., Kondo H., Yokoi T.,
Kodaira H., Furuya T., Takahashi M., Kikkawa E., Omura Y., Abe K.,
Kamihara K., Katsuta N., Sato K., Tanikawa M., Yamazaki M.,
Sugiyama T., Irie R., Otsuki T., Sato H., Wakamatsu A., Ishii S.,
Yamamoto J., Isono Y., Kawai-Hio Y., Saito K., Nishikawa T.,
Kimura K., Yamashita H., Matsuo K., Nakamura Y., Sekine M.,
Kikuchi H., Murakawa K., Kanehori K., Takahashi-Fujii A.,
Oshima A., Sugiyama A., Kawakami B., Suzuki Y., Sugano S.,
Nagahari K., Maeuho Y., Nagai K. and Isogai T.
NEO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4136)
AUTHORS Isogai T. and Yamamoto J.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kanagari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5' & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .4136

/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE3002390"
/tissue_type="cerebellum"
/clone_lib="BRACE3"
/notes="Cloning vector: pME18SFL3"
136..537
/notes="unnamed protein product"
/codon_start=1
/db_xref="GI:34530325"
/translation="MVMLEDSNSTGCGARNCECLVFLSVLGCOSRKGMQRTOQAG
RWLRAGREASSETNPGTLLDQSPPLPLAARGWQHEHPVRCNVLPHAFSSNCFQG
NPFKWSQTQLNRVTYKRWELLSCNSGRK"

CDS
Alignment Scores: 4.3e-51 Length: 4136
Pred. No.: 581.00 Matches: 114
Score: 581.00 Mismatches: 1
Percent Similarity: 85.19% Conservative: 1
Best Local Similarity: 84.44% Mismatches: 0
Query Match: 81.26% Indels: 20
DB: 9 Gaps: 1
US-10-005-907-2 (1-135) x AK124520 (1-4136)
QY 1 MetGlyAsnTyrLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 498 ATGGGAATATCTCTCGCAAACTC----- 524
QY 21 LysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 525 -----AGCAGGAATGACTACATTTCAGAAACTT 557
QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
DB 558 CAAGATCAAGATAAGAAAGCAAGAGTTTCACACTTCTAATCAGGAAACGAGAAT 617
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 618 GGCAAGTGGTCTGGAAGAGTGCTGACACTGTCATTATCATATCCCATCCCAAGAAATCC 677
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 678 TCCTGAGTCCCATGATGATGGCTATGAGACATTCACCTCCCAAGAAAGTGAGA 737
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 738 CAGTTTAGAAGAAAGGTGAGACAGAGATATGCCCTTCTTAGGACTTCTGTAGAGCCT 797
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 798 TGTTCCTGCACCATGAGCATGATTATGAAGTTGTGTTCACAC 842

RESULT 7
AC074365/c 141268 bp DNA linear HTG 23-SEP-2000
LOCUS Homo sapiens chromosome 1 clone RF11-115C4, WORKING DRAFT SEQUENCE,
10 unordered pieces.
ACCESSION AC074365
VERSION AC074365.5 GI:10280935
KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 141268)
AUTHORS Waterston, R.H.
TITLE The sequence of Homo sapiens clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 141268)
AUTHORS Waterston, R.H.

```


TITLE Direct Submission
JOURNAL Submitted (29-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
COMMENT On Sep 23, 2000 this sequence version replaced gi:9838075.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site:http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH01SC04
----- Summary Statistics -----
Sequencing vector: M13, 100%
Chemistry: plasmid;
Chemistry: Dye-primed ET; 100% of reads
Chemistry: Dye-terminator Big Dye; 0% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 136274 bases at least Q40
Consensus quality: 138356 bases at least Q30
Consensus quality: 139182 bases at least Q20
Insert size: 142000; agarose-fp
Insert size: 140368; sum-of-contigs
Quality coverage: 4.64 in Q20 bases; agarose-fp
Quality coverage: 4.81 in Q20 bases; sum-of-contigs

```

* NOTE: this is a 'working draft' sequence. It currently
* contains only contigs where the difference between the
* adjacent known contigs or the difference between the
* adjacent unknown contigs are represented as
* 'N'. But the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence.
* As soon as it is available and the accession number will
* be preserved.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45	46	47	48	49	50	51	52	53	54	55	56	57	58	59	60	61	62	63	64	65	66	67	68	69	70	71	72	73	74	75	76	77	78	79	80	81	82	83	84	85	86	87	88	89	90	91	92	93	94	95	96	97	98	99	100	101	102	103	104	105	106	107	108	109	110	111	112	113	114	115	116	117	118	119	120	121	122	123	124	125	126	127	128	129	130	131	132	133	134	135	136	137	138	139	140	141	142	143	144	145	146	147	148	149	150	151	152	153	154	155	156	157	158	159	160	161	162	163	164	165	166	167	168	169	170	171	172	173	174	175	176	177	178	179	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196	197	198	199	200	201	202	203	204	205	206	207	208	209	210	211	212	213	214	215	216	217	218	219	220	221	222	223	224	225	226	227	228	229	230	231	232	233	234	235	236	237	238	239	240	241	242	243	244	245	246	247	248	249	250	251	252	253	254	255	256	257	258	259	260	261	262	263	264	265	266	267	268	269	270	271	272	273	274	275	276	277	278	279	280	281	282	283	284	285	286	287	288	289	290	291	292	293	294	295	296	297	298	299	300	301	302	303	304	305	306	307	308	309	310	311	312	313	314	315	316	317	318	319	320	321	322	323	324	325	326	327	328	329	330	331	332	333	334	335	336	337	338	339	340	341	342	343	344	345	346	347	348	349	350	351	352	353	354	355	356	357	358	359	360	361	362	363	364	365	366	367	368	369	370	371	372	373	374	375	376	377	378	379	380	381	382	383	384	385	386	387	388	389	390	391	392	393	394	395	396	397	398	399	400	401	402	403	404	405	406	407	408	409	410	411	412	413	414	415	416	417	418	419	420	421	422	423	424	425	426	427	428	429	430	431	432	433	434	435	436	437	438	439	440	441	442	443	444	445	446	447	448	449	450	451	452	453	454	455	456	457	458	459	460	461	462	463	464	465	466	467	468	469	470	471	472	473	474	475	476	477	478	479	480	481	482	483	484	485	486	487	488	489	490	491	492	493	494	495	496	497	498	499	500	501	502	503	504	505	506	507	508	509	510	511	512	513	514	515	516	517	518	519	520	521	522	523	52
--	---	---	---	---	---	---	---	---	---	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	----

```

FEATURES
  source
    1. 141268
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /chromosome="1"
      /clone="RP11-115C4"
    1. 31377
      /note="assembly_name:Contig10"
    31478. 63011
      /note="assembly_name:Contig11"
    63121. 64845
      /note="assembly_name:Contig2"
    64946. 67606
      /note="assembly_name:Contig3"
    67707. 77674
      /note="assembly_name:Contig4"
    77775. 89256
      /note="assembly_name:Contig5"
    89357. 100400
      /note="assembly_name:Contig6"

```

```
misc_feature      /note="assembly_name:Contig6"
10504..11915
/note="assembly_name:Contig7"
misc_feature      112016..125713
/note="assembly_name:Contig8"
125814..141268
/note="assembly_name:Contig9
clone end:SP6
vector side:left"
```

ORIGIN

```

Alignment Scores:
Pred. NO.: 1.58e-33
Score: 430.00
Percent Similarity: 100.00%
Best Local Similarity: 100.00%
Query Match: 60.14%
DB: 2
Length: 141268
Matches: 80
Conservative: 0
Mismatch: 0
Indels: 0
Gaps: 0

```

TIS-10-005-907-2 (1-135) x AC074365 (1-141268)

QY	56	GlnGluAsnGluAenGlySerGlySerGluGluValCysTyrThrValIleAsnHisIle	75
Db	5700	CAGAAAAACGAGAATGCAGTGCTTCGTGAAGAAGTGTGTACTCACTGTATTAAATCACATC	5641
QY	76	ProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAsnSerLeu	95
Db	5640	CCCCATCAGAGATCCTCCCTGGAGCTCAATGATGATGGCTATGAGAAAATTGACTGCCCTC	5581
QY	96	ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr	115
Db	5580	ACNAGGAAGTGACACAGTTTAGAANAAGGTGNAGACACAATATGCCCTCTTAGGACT	5521
QY	116	SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis	135
Db	5520	TCTGTTACTAGGCTTCTCTCGACCCCATGACCATGATTATGAAGTTGTGTTCCACAC	5461

RESULT 8						PRI 23-OCT-2002
AL606804						
LOCUS	185467 bp	DNA	linear			
DEFINITION	Human DNA sequence from clone RP11-978I15					
						on chromosome 1, complete sequence.

AL606804
AL606804.11
HTG

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 185467)

AUTHORS	Almeida, J.
TITLE	Direct Sub
JOURNAL	Submitted

Cambridgehire, CB10 1SA, UK. E-mail enquiries: humuery@sanger.ac.uk
 On Oct 24, 2002 this sequence version replaced gi.18121511.
 During sequence assembly data is compared from overlapping clones.
 When differences are found these are annotated as variations
 together with a note of the overlapping clone name. Note that the
 variation annotation may not be found in the sequence submission
 corresponding to the overlapping clone, as we submit sequences with
 only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Emi, ENMBL; Sw., SWISSPROT; Tr., TREMBL; Wp., WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at <http://www.sanger.ac.uk/HGP/Chr1>. RP11-978115 is from the library RP11-11.4 constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBACe3.6
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

FEATURES

source
1. 185467
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="1"
/clone="RP11-978115"
/clone_lib="RP11-11.4"

ORIGIN

Alignment Scores:
Pred. No.: 2.13e-33 Length: 185467
Score: 430.00 Matches: 80
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 60.14% Indels: 0
DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x AL606804 (1-185467)

Qy 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrVallieAanHisile 75
Db 132080 CAGGAAACAGGAATGCGAGTGGTTCGGAAGAGTGTGTACACTGTCAATATCATCATC 132139
Qy 76 ProHisGlnArgSerLeuSerSerAsnAspGlyTyrGluAsnLeuAspSerLeu 95
Db 132140 CCCCATCAGAGATCCCTCCCTGAGCTCAATGATGATGCTATGAGACATGACTCCCTC 132199
Qy 96 ThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThr 115
Db 132200 ACAAGAAAGTGACAGACTTTAGGAAGGTGACAGACAGAAATGCGCTCTTAGGACT 132259
Qy 116 SerValSerArgProCysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
Db 132260 TCTGTAGTAGGCTTGTCTGTCACCCATGAGCATGATTATGAAGTTGTGTTTCCACAC 132319

RESULT 9
BC030506 1332 bp mRNA linear PRI 06-OCT-2003
LOCUS Homo sapiens germinal center expressed transcript 2, mRNA (CDNA
clone MGC:40441 IMAGE:4385178), complete cds.
VERSION BC030506
ACCESSION BC030506
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
1 (bases 1 to 1532)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS
Strausberg, R.D., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenon, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Haieh, F.,
Diatchenko, L., Narusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Sapich, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Schaeetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abramson, R.D., Mullany, S.J., Bosak, S.A., McEwan, P.J.,

McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W.,
Villalón, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahey, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalios, D.E.,
Schnerch, A., Schein, J.E., Jones, S.J., and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
22388357
12477932

2 (bases 1 to 1532)

Strausberg, R.

Direct Submission

Submitted (07-MAY-2002) National Institutes of Health, Mammalian

Gene Collection (MGC), Cancer Genomics Office, National Cancer

Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,

USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgabbs-remail.nih.gov

Tissue Procurement: Lou Staudt

CDNA Library Preparation: Life Technologies, Inc.

CDNA Library Arrayed by: the I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: National Institutes of Health Intramural

Sequencing Center (NISC),

Gaithersburg, Maryland;

Web site: <http://www.nisc.nih.gov/>

Contact: nisc.mgc@nih.gov

Akhter, N., Ayala, K., Beckstrom-Sternberg, S.M., Benjamin, B.,

Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S.,

Dietrich, N.L., Granito, S., Guan, X., Gupta, J., Hagnigni, P.,

Hansen, N., Ho, S.-L., Karlins, S., Kwong, P., Laric, P., Legaspi, R.,

Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,

McDowell, J., Pearson, R., Stantirpop, S., Thomas, P.J., Touchman, J.W.,

Tsurgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,

Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
Series: IRAK Plate: 64 Row: a Column: 15

This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 22749536.

FEATURES

source

1. 1532
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="MGC:40441 IMAGE:4385178"
/tissue_type="Lymph, lymphoma"
/clone_lib="NIH MGC_85"
/lab_host="DH10B"
/note="Vector: pCMV-SPORT6"
1. 1532
/gene="GCET2"
/note="synonyms: HGAL, GCAT2, MGC40441"
/db_xref="LocusID:257144"
/db_xref="MIM:607792"

gene

1. 1532

/product="germinal center expressed transcript 2"
/protein_id="AAH30506.1"
/db_xref="GI:20987806"
/translation="WGNLSLENRQQTQEMPNVNRMQSPKORTSCWDHHAEGCF
CLPWKILIFKRODSQNERMSSTPIQNDVDTYSELCTYLNEHVLCTRPSGNS
AEEYENVPCKAARPSLGGTETETISLHMPSTDFRPARSPFEYELLMPHRISSHF
LQQRPLMAPSETQFSLH"

CDS

54..590

/codon_start=1

/product="germinal center expressed transcript 2"

/protein_id="AAH30506.1"

/db_xref="GI:20987806"

/translation="WGNLSLENRQQTQEMPNVNRMQSPKORTSCWDHHAEGCF
CLPWKILIFKRODSQNERMSSTPIQNDVDTYSELCTYLNEHVLCTRPSGNS
AEEYENVPCKAARPSLGGTETETISLHMPSTDFRPARSPFEYELLMPHRISSHF
LQQRPLMAPSETQFSLH"

ORIGIN

Alignment Scores: 2.76e-08 Length: 1532
Pred. No.: 169.50 Matches: 47
Score: 51.61% Conservative: 33
Percent Similarity: 30.32% Mismatches: 54
Best Local Similarity: 23.71% Indels: 21
Query Match: 9 Gaps: 5
DB: 5

US-10-005-907-2 (1-135) x BC030506 (1-1532)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 54 ATGGGAATCTCTCTGAGAGAAACAGCGCGCAGACACTCAAGAGATCGCTGG 113
QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 114 AATGTGAGATGCAAAAGCCCAAGACAGCAATCCAGATCGTGGATCACCATATCGCT 173
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGluAsp 42
Db 174 GAAGGGTGTCTGCTCCATCGGAAATAATCACTATTTTGAAGAGG---CAAGAT 230
QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlySer 62
Db 231 TCCCAAAACGAAATCAAGAGTACTATCCATCCAGCAATGTTGACAGACC 290
QY 63 GlySerGluGluValCysThrValIleAsnHis---IleProHisGlnArgSer 81
Db 291 TACTCAGAGGAGCTGTCTATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 350
QY 82 LeuSerAsnAspAspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArgGln 101
Db 351 GGGAACTCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGAGCTGAGAGACCAGAG 410
QY 102 PheArgGluArgSerGluThrValAlaLeuArgLysValSerValSerArgProCys 121
Db 411 TCCTTGGAGGAGACTGAGTATGATTCATCATATATGCTTCTTACAGACCCAGG 470
QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 471 CATGCCCGATCCCGAGAGATGAATATGAATCTTCATCGCTCAC 515

RESULT 10
AF521911 1654 bp mRNA linear PRI 14-JAN-2003
LOCUS
DEFINITION Homo sapiens HGAL mRNA, complete cds.
ACCESSION AF521911
VERSION AF521911.1 GI:27733682
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1654)
Losos, I.S., Alizadeh, A., Rajapaksa, R., Tibshirani, R. and Levy, R.
HGAL is a novel interleukin-4-inducible gene that strongly predicts survival in diffuse large B-cell lymphoma
Blood 101 (2), 433-440 (2003)
JOURNAL
MEDLINE 22397484
PUBMED 12509382
REFERENCE 2 (bases 1 to 1654)
Losos, I.S., Alizadeh, A.A. and Levy, R.
Direct Submission
TITLES
SUBMITTED (16-JUN-2002) Division of Oncology, Department of Medicine, Stanford University, 269 Campus Drive, CCSR Building, Room 1100, Palo Alto, CA 94305-5151, USA
JOURNAL
FEATURES
source
1. 1654
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
173..709
/note="germinal center-associated lymphoma; contains

modified ITAM motif, similar to mouse M17"
/codon_start=1
/product="HGAL"
/protein_id="AA021701.1"
/db_xref="GI:27733683"
/translation="MNSLLRENRQONTQEMVNRMSKORTSRQWDHIIARGCF
ALPEWKILIFKRDQSNENRMSSTPIQDAVDQTYSELCLYLNHSLVLCITRPSGNS
ABEYENVPCKAERPRESLGGTETETISLHMPSTDFRHPARSPEDYELLMHPRISSHF
LQOPRLMAPSETQFSLH"

ORIGIN
Alignment Scores: 3e-08 Length: 1654
Pred. No.: 169.50 Matches: 47
Score: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 5

US-10-005-907-2 (1-135) x AF521911 (1-1654)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 173 ATGGGAATCTCTCTGAGAGAAACAGCGCGCAGACACTCAAGAGATCGCTGG 232
QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 233 AATGTGAGATGCAAAAGCCCAAGACAGCAATCCAGATCGTGGATCACCATATCGCT 292
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGluAsp 42
Db 293 GAAGGGTGTCTGCTCCATCGGAAATAATCACTATTTTGAAGAGG---CAAGAT 349
QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlySer 62
Db 350 TCCCAAAACGAAATCAAGAGTACTATCCATCCAGCAATGTTGACAGACCAGAG 409
QY 63 GlySerGluGluValCysThrValIleAsnHis---IleProHisGlnArgSer 81
Db 410 TACTCAGAGGAGCTGTCTATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 469
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnLeuAspSerLeuThrArgLysValArgGln 101
Db 470 GGGAACTCTGCTGAAGAGTACTATGAGAATGTTCCCTGCAAGAGCTGAGAGACCAGAG 529
QY 102 PheArgGluArgSerGluThrValAlaLeuArgLysValSerValSerArgProCys 121
Db 530 TCCTTGGAGGAGACTGAGTATGATTCATCATATATGCTTCTTACAGACCCAGG 589
QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 590 CATGCCCGATCCCGAGAGATGAATATGAATCTTCATCGCTCAC 634

RESULT 11
AF5212246 3270 bp mRNA linear PRI 14-OCT-2003
LOCUS
DEFINITION Homo sapiens germinal center B-cell expressed transcript 2 (GCET2)
mRNA, complete cds.
ACCESSION AF5212246
VERSION AF5212246.1 GI:27948576
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 3270)
Pan, Z., Shen, Y., Du, C., Zhou, G., Rosenwald, A., Staudt, L.M., Greiner, T.C., McKeithan, T.W. and Chan, W.C.
Two newly characterized germinal center B-cell-associated genes, GCET1 and GCET2, have differential expression in normal and neoplastic B cells
Am. J. Pathol. 163 (1), 135-144 (2003)
JOURNAL 22702315
MEDLINE

PUBMED 12819018
 REFERENCE 2 (bases 1 to 3270)
 AUTHORS Pan,Z., Shen,Y., Du,C., Zhou,G., McKeithan,T., Rosenwald,A.,
 Staudt,J., and Chan,W.C.
 TITLE Direct Submission
 JOURNAL Submitted (07-JAN-2003) Pathology, University of Nebraska Medical
 Center, #42, Omaha, NE 68198, USA
 FEATURES
 source
 1..3270
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /map="3q13.13"
 /cell_line="DHL16"
 /gene="GCBT2"
 /notes="GCBT2"
 /notes="GCBT2"
 107..643
 /gene="GCBT2"
 /notes="similar to Mus musculus germinal center transcript
 M17; similar to MGC40441 sequence in GenBank Accession.
 Number BC030506"
 /codon_start=1
 /product="germinal center B-cell expressed transcript 2"
 /protein_id="AA022147.1"
 /db_xref="GI:27948577"
 /translations="MGNLSLRNRQNTQEMPNVRMQSPKORTSRCDWHIAEGCF
 CLPWKILIEKQDSNENRMSSTPIQNVDTQYSEELCYTLINRHLVLCRPSGNS
 AEEYVNPCKARPRESLGTETVSYLLHMPSTDPRHARSPEDEYELLMPHRISSHF
 LQPRPLMAPSERQFSLH"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,34e-08 Length: 3270
 Score: 169.50 Matches: 47
 Percent Similarity: 51.61% Conservatives: 33
 Best Local Similarity: 30.32% Mismatches: 54
 Query Match: 23.71% Indels: 21
 DB: 9 Gaps: 5
 US-10-005-907-2 (1-135) x AY212246 (1-3270)
 QY 1 MetGlyAsnTyrLeuLeuArgGlyLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
 Db 107 ATGGGAATCTCTGCTGAGAGAAACAGCGGCGAGCAGACACTCAAGAGATGCTTGG 166
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
 Db 167 AACGTGAGATGCAAGAGCCCAACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 226
 QY 31 -----GlnGluMetThrThrPheGluArgGlyLeuGlnAsp 42
 Db 227 GAAGGGTGTCTGCTTCCATGAGAAATAATCTCATTTTGAAGAGAG---CAAGAT 283
 QY 43 GlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
 Db 284 TCCCAACGAAATGAAGAATGTCATCTACTCCCATCCAGGACAATGTTGACCAACC 343
 QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 81
 Db 344 TACTCAGAGGAGGTGTGCTATACCTCATCAATCATCGGTCTCTGTACAGGCCATCA 403
 QY 82 LeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
 Db 404 GGGAACTCTGTGAAGACTACTATGAGATGTTCTCCGCAAGCTGAGAGCCCAAGAG 463
 QY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
 Db 464 TCCITGGGAGGAACCTGAGACTGAGTATTCACTTCTACATATGCTTCTACAGCCCCAGG 523
 QY 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135

Db 524 CATGCCCGATCCCCAGAGAGATGAATATGAACCTTCTCATGCTCAC 568
 RESULT 12
 AX740457
 LOCUS 1643 bp DNA linear PAT 10-MAY-2003
 DEFINITION Sequence 46 from Patent WO02020756.
 AX740457
 ACCESSION AX740457.1 GI:30523621
 VERSION
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Chordata; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 AUTHORS
 1 Stuart,J., Lincoln,S.E., Altus,C.M., Dufour,G., Chalup,M.,
 Hillman,J.L., Jones,A.L., Yu,J.Y., Wright,R.J., Gietzen,D., Liu,T.,
 Yap,P., Dahl,C.R., Momiyama,M.G., Bradley,D., Rohatgi,S.,
 Harris,B., Roseberry, Ann.M., Gerstin,E.H., Peralta,C.H.,
 David,M.H., Panzer, Scott.R., Flores,V., Daffo,A., Marwaha,R.,
 Chen,A.J., Chang,S.C., Au,A.P. and Inman,R.R.
 TITLE Secretory molecules
 JOURNAL Patent: WO 02020756-A 46 14-MAR-2002;
 Incyte Genomics, Inc. (US)
 FEATURES
 source
 1..1643
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /note="Incyte ID No: LG:018739.2:2000SEP08"
 ORIGIN
 Alignment Scores:
 Pred. No.: 3,79e-07 Length: 1643
 Score: 159.00 Matches: 47
 Percent Similarity: 50.64% Conservatives: 22
 Best Local Similarity: 30.13% Mismatches: 55
 Query Match: 22.24% Indels: 22
 DB: 6 Gaps: 6
 US-10-005-907-2 (1-135) x AX740457 (1-1643)
 QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysPro--- 19
 Db 1096 ATGGGAATCTCTGCTGAGAGAAACAGCGGCGAGCAGACACTCAAGAGATGCTTGG 1155
 QY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
 Db 1156 ATGTGAGATCAAGAGCCCAACAGAGAACATCCAGATGCTGGGATCACCATATCGCT 1215
 QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
 Db 1216 GAAGGGTGTCTGCTTCCATGAGAAATAATCTCATTTTGAAGAGAG---CAAGAT 1272
 QY 43 GlnAspLysLysSerGlnGluValSerSerThrSer---AsnGlnGluAsnGly 61
 Db 1273 TCCCAACGAAATGAAGAATGTCATCTACTCCCATCCAGCAGACATGTTGACCAG 1332
 QY 62 SerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSer 80
 Db 1333 ACCTACTCAGAGAGGTGTGCTATACCTCATCAATCATCGGTCTCTGTACAGGCCA 1392
 QY 81 SerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 Db 1393 TCAGGAATCTGCTGAAGAGTACTATGAGATGTTCCCTCGAAGCTGAGAGCCCA 1452
 QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 Db 1453 GAGTCTCTGGGAGGAACCTGAGACTGAGTATTCACTTCTANATATGCTTCTACAGCCCC 1512
 QY 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
 Db 1513 AGGCATGCCCGATCCCCAGAGAGATGAATATGAACCTTTCATGCCTCAC 1560

```

RESULT 13
MMU13263
LOCUS Mus musculus M17 mRNA, complete cds. 1672 bp mRNA linear ROD 15-APR-1996
DEFINITION Mus musculus M17 mRNA, complete cds.
ACCESSION U13263
VERSION U13263.1 GI:533130
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1672)
AUTHORS Christoph, T., Rickert, R. and Rajewsky, K.
TITLE M17: a novel gene expressed in germinal centers
JOURNAL Int. Immunol. 6 (8), 1203-1211 (1994)
PUBMED 95071950
REFERENCE 2 (bases 1 to 1672)
AUTHORS Rickert, R.C.
TITLE Direct Submission
JOURNAL Submitted (11-AUG-1994) Robert C. Rickert, Institute for Genetics, University of Cologne, Weyertal 121, Cologne, 50823, Germany
FEATURES
source
1. 1672
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6"
/db_xref="taxon:10090"
gene
1. 1672
/gene="M17"
CDS
94..573
/gene="M17"
/codon_start=1
/product="M17 protein"
/protein_id="AA096821.1"
/db_xref="GI:1262852"
/translation="MGNCILQRTYFRGMSCCSVSCCLPWNRTFTKARQESPKQ
NEGTSAPVDNANETTELCLYLDVHEAVRPSVNPABGYENISKAERHKSS
RGTEYSLVRFPPQPLPSTDDYELLPSRFSHFQOPLPTTPTVTHFSYQ"
585..727
/gene="M17"
3' UTR
/gene="M17"
ORIGIN
Source
1. 4218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_clone="PLACE7001544"
/tissue_type="placenta"
/clone_lib="PLACE7"
/note="cloning vector: pME18SPL3"
ORIGIN
Alignment Scores:
Pred. No.: 0.00315 Length: 4218
Score: 126.00 Matches: 30
Percent Similarity: 59.76% Conservative: 19
Best Local Similarity: 36.59% Mismatches: 31
Query Match: 17.62% Indels: 2
Gaps: 2
US-10-005-907-2 (1-135) x AK125521 (1-4218)
QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal 50
Db 181 AAAACATACGCACATTTAAAGCCAGA---CAAGAGTCTCCAAAGCAAAATGAGGATG 237
QY 51 SerSerThrSerAsnGlnAsnGlnAsnGlySerGlySerGluGluValCysTyrThr 70
Db 238 ACTTCAGCTCCCGTTTCAGCAACATGCTAATGAGACCTACACAGAGAGTGTGTCTATC 297
QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspGlyTyr 89
Db 298 CTGTGTGATCAGAGCTGTCAAGGGAAGGCCATCAGTGAACCTCTGAGGGGTTCTAC 357
QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
Db 358 CAGACATCTCTTAAACAGCTGACAGACACAAAGAGCTCTTCAAGAGGGAACAGAGACTGAG 417
QY 110 TyrAlaLeuLeuArg---ThrSerValSerArgProCysSerCysThrHisGluHisAsp 128
Db 418 TATTCGGTTCCTCCGTTTCCTCTTCCCTCCTCAGCCCTACCTTCCACA---GATGATGAA 474

```

```

QY 129 TyrGluValValPhePro 134
Db 475 TATGAACCTCTTATGCCC 492
RESULT 14
AK125521/c
LOCUS Homo sapiens cDNA FLJ43533 fis, clone PLACE7001544.
DEFINITION Homo sapiens cDNA FLJ43533 fis, clone PLACE7001544.
ACCESSION AK125521
VERSION AK125521.1 GI:34531645
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1
AUTHORS Tashiro, H., Yamazaki, M., Watanabe, K., Kumagai, A., Itakura, S., Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K., Watanabe, M., Murakawa, K., Kanetori, K., Takahashi-Fujii, A., Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S., Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
TITLE NEDO human cDNA sequencing project
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 4218)
AUTHORS Isogai, T. and Yamamoto, J.
TITLE Direct Submission
JOURNAL Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7 Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan (E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
COMMENT NEDO human cDNA sequencing project supported by Ministry of Economy, Trade and Industry of Japan; cDNA full insert sequencing: Research Association for Biotechnology (RAB); cDNA library construction: Helix Research Institute (HRI) (supported by Japan Key Technology Center etc.); 5' & 3' end one pass sequencing: RAB, HRI, and Biotechnology Center, National Institute of Technology and Evaluation; clone selection for full insert sequencing: HRI and RAB; annotation: HRI and RAB.
FEATURES
Source
1. 4218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/db_clone="PLACE7001544"
/tissue_type="placenta"
/clone_lib="PLACE7"
/note="cloning vector: pME18SPL3"
ORIGIN
Alignment Scores:
Pred. No.: 0.00315 Length: 4218
Score: 126.00 Matches: 30
Percent Similarity: 59.76% Conservative: 19
Best Local Similarity: 36.59% Mismatches: 31
Query Match: 17.62% Indels: 2
Gaps: 2
US-10-005-907-2 (1-135) x AK125521 (1-4218)
QY 56 GlnGluAsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHis--- 74
Db 3384 CAGGACAATGTTGACCAGCAGCTTCTCAGGAGCTGTGTCTATCCTCATCATCGG 3325
QY 75 IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSer 94
Db 3324 GTTCTCTGTACAGGCCATTCAGGGAACCTCTGCTGAAGAGTACTATGAGAATGTTCCCTGC 3265
QY 95 LeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluThrAlaLeuLeuArg 114
Db 3264 AAGCTGAGAGAGCCAGAGAGTCTTGGAGGAACTGAGAGCTGAGTATTCTTCTCAT 3205

```

QY 115 ThrSerValSerArgProCysSerCys---ThrHisGluHisAspTyrGluValValPhe 133
 Db 3204 ATGCTTCTACACACCCCGATGCGCCGATCCCGAGAGATGATATGACTTTCATG 3145
 QY 134 ProHis 135
 Db 3144 CCTCAC 3139
 RESULT 15
 AC128688
 LOCUS Homo sapiens 3 BAC RP11-757F18 (Roswell Park Cancer Institute Human
 DEFINITION BAC Library) complete sequence.
 AC128688
 VERSION AC128688.4 GI:22549624
 KEYWORDS HTG.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 81001)
 AUTHORS Muzny,D.M., Adams,C., Adio-Oduola,B., Ali-osman,F.R., Allen,C.,
 Alsbrooks,S.L., Anarunge,H.C., Are,J.R., Ayele,M., Banks,T.,
 Barbara,J., Benton,J., Benge,K., Blankenburg,K., Bonnin,D.,
 Bouck,J., Bowie,S., Brieva,M., Brown,E., Brown,N., Bryant,N.P.,
 Buhay,C., Burch,P., Burkett,C., Burrell,K.L., Byrd,N.C.,
 Carron,T.F., Carter,M., Cavazos,S.R., Chacko,J., Chavez,D.,
 Chen,G., Chen,R., Chen,Z., Chowdhry,I., Christopoulos,C.,
 Cleveland,C.D., Cox,C., Coyle,M.D., Dathorne,S.R., David,R.,
 Davila,M.L., Davis,C., Davy-Carroll,L., Dederich,D.A.,
 Delaney,K.R., Delgado,O., Denn,A.L., Ding,Y., Dinh,H.H.,
 Douthwaite,K.J., Draper,H., Dugan-Rocha,S., Durbin,K.J.,
 Earnhart,C., Edgar,D., Edwards,C.C., Elhaj,C., Escotto,M.,
 Falle,T., Ferraguto,D., Flagg,N., Ford,J., Foster,P., Frantz,P.,
 Gabisi,A., Gao,J., Garcia,A., Garner,T., Garza,N., Gill,R.,
 Gorrill,J.H., Guevara,W., Gunaratne,P., Hale,S., Hamilton,K.,
 Harris,C., Harris,K., Hart,M., Havlak,P., Hawes,A., He,X.,
 Hernandez,J., Hernandez,O., Hodgson,A., Hognes,M., Holloway,C.,
 Hollins,B., Honsi,F., Howard,S., Huber,J., Hulyk,S., Hume,J.,
 Jackson,L.E., Jacobson,B., Jia,X., Johnson,R., Jolivet,S.,
 Joudan,S., Karlsson,E., Kelly,S., Khan,U., King,L., Korvah,J.,
 Kovar,C., Kratovic,J., Kureshi,A., Landry,N., Leal,B., Lewis,L.C.,
 Lewis,L., Li,J., Li,Z., Lichtarge,O., Lieu,C., Liu,J., Liu,W.,
 Lougheed,H., Lozado,R.J., Lu,X., Lucier,A., Lucier,R., Luna,R.,
 Ma,J., Maheshwari,M., Mapua,P., Martin,R., Martindale,A.,
 Martinez,E., Massey,E., Mawhney,E., McLeod,M.P., Meador,M.,
 Mei,G., Metzker,M., Miner,G., Miner,Z., Mitchell,T., Mohabbat,K.,
 Moore,S., Morgan,M., Moorish,T., Morris,S., Moser,M., Neal,D.,
 Nelson,D., Newton,J., Newton,N., Nguyen,A., Nguyen,N., Nguyen,N.,
 Nickerson,E., Nwokenkwo,S., Oguh,M., Okwuonu,G., Oragunye,N.,
 Oviedo,R., Pace,A., Payton,B., Peery,J., Perez,L., Peters,L.,
 Pickens,R., Primus,E., Pu,L.L., Quiles,M., Ren,Y., Rives,M.,
 Rojas,A., Rojibokan,I., Rolfe,M., Ruiz,S., Saverly,G., Scherer,S.,
 Scott,G., Shen,H., Shoohtari,N., Sisson,I., Sodergren,E.,
 Sonake,T., Sparks,A., Stanley,H., Stone,H., Sutton,A., Svatek,A.,
 Taylor,P., Tamerisa,A., Tamerisa,K., Tang,H., Tansey,J., Taylor,C.,
 Taylor,T., Telford,B., Thomas,N., Thomas,S., Usmani,K., Vasquez,L.,
 Vera,V., Villalón,D., Vinson,R., Wang,Q., Wang,S., Ward-Moore,S.,
 Warren,R., Washington,C., Watlington,S., Williams,G.,
 Williamson,A., Wlezyk,R., Wooden,S., Worley,K., Wu,C., Wu,Y.,
 Wu,Y.F., Zhou,J., Zorrilla,S., Naylor,S.L., Weinstein,G. and
 Gibbs,R.
 Direct Submission
 Unpublished
 REFERENCE 2 (bases 1 to 81001)
 AUTHORS Worley,K.C.
 TITLE Direct Submission
 JOURNAL Submitted (22-JUL-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 REFERENCE 3 (bases 1 to 81001)
 AUTHORS Worley,K.C.
 TITLE Direct Submission

JOURNAL

REFERENCE
 AUTHORS
 TITLE
 JOURNAL

COMMENT

Submitted (26-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 4 (bases 1 to 81001)
 Worley,K.C.
 Direct Submission
 Submitted (30-AUG-2002) Human Genome Sequencing Center, Department
 of Molecular and Human Genetics, Baylor College of Medicine, One
 Baylor Plaza, Houston, TX 77030, USA
 On Aug 30, 2002 this sequence version replaced gi:22475264.
 INFORMATION: <http://www.hgsc.bcm.tmc.edu/> or email
gc-help@bcm.tmc.edu

CLONE LENGTH: This sequence does not necessarily represent the
 entire insert of this clone. Overlapping regions of clones are only
 sequenced and submitted once, so the sequence for the remainder of
 the insert may be found in the record for the adjacent clones.
 Overlapping clones are noted at the beginning and end of the
 Features listing.

ANNOTATION OF FEATURES:

STSs are identified using ePCR (Genome Res. 7:541-550) searches
 of a local database that includes entries from dbSTS, GDB, and
 local mapping efforts.

Repeats are identified using RepeatMasker (A. Smit and P. Green,
 unpublished.) for Human and Mouse sequences.
 Genes and Region of sequence similarity are identified by BLAST
 (Nuc. Acids Res. 25:3389-3402) similarity (expect < 1e-34) to the
 EST and cDNA sequences. Genes demonstrate at least two exons
 flanked by consensus splice sites that maintained sequence
 continuity across the splice junctions. Sequences that are not
 identical matches are annotated as similar.

SEQUENCING READ COVERAGE: Sequencing is completed to a minimum
 standard of double strand coverage with a minimum of 2 clones and 2
 reads with no ambiguities or 2 chemistries with a minimum of 2
 clones and 3 reads with no ambiguities. If the sequence quality for
 a region does not meet this standard, it will be indicated in the
 annotation as Low Coverage.

QUALITY OF INDIVIDUAL BASES: This sequence meets stringent quality
 standards - estimated error rate less than 1 per 10,000 bases.
 Reports of lowest quality individual bases and measures of base
 quality are listed below. Description of the metrics can be found
 at URL:
<http://www.hgsc.bcm.tmc.edu:8088/quality.info/genbank.annotation.ht>

FEATURES

Location/Qualifiers

1. .81001
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /chromosome="3"
 /clone="RP11-757F18"
 complement(1. .2041)
 /note="overlaps bases 1. .2041 of clone AC119734"
 /function="clone overlap"
 11149. .11300
 /standard_name="RH65506"
 27049. .27072
 /note="Size confirmed by PCR and restriction digest."
 /function="pcr product sequence only"
 31707. .31920
 /standard_name="G44369"
 32757. .33036
 /standard_name="D1183206"
 68551. .68832
 /standard_name="SHGC-112515"
 71827. .71979
 /standard_name="RH99175"
 74784. .75043
 /standard_name="SHGC-68798"
 77907. .78053

misc_feature

STS

misc_feature

STS

STS

STS

STS

STS

STS

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 11:26:15 ; Search time 1784 Seconds
(without alignments)
9839.646 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429
Perfect score: 405
Sequence: 1 atgggaattatctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb.ba.*

2: gb.rtg.*

3: gb.in.*

4: gb.om.*

5: gb.ov.*

6: gb.pat.*

7: gb.ph.*

8: gb.pl.*

9: gb.pr.*

10: gb.ro.*

11: gb.sts.*

12: gb.sy.*

13: gb.un.*

14: gb.vi.*

15: em.ba.*

16: em.fun.*

17: em.hum.*

18: em.in.*

19: em.mu.*

20: em.or.*

21: em.ov.*

22: em.pat.*

23: em.ph.*

24: em.pl.*

25: em.ro.*

26: em.sts.*

27: em.un.*

28: em.vi.*

29: em.vrt.*

30: em.htg.hum.*

31: em.htg.inv.*

32: em.htg.other.*

33: em.htg.mus.*

34: em.htg.pln.*

35: em.htg.rod.*

36: em.htg.mam.*

37: em.htg.vrt.*

38: em.sy.*

39: em.htgo.hum.*

40: em.htgo.mus.*

41: em.htgo.other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	405	100.0	2648	9	BC024174	Homo sapi
2	405	100.0	3762	6	AX505122	Sequence
3	403.4	99.6	1709	9	AK123798	Homo sapi
4	359.4	88.7	1924	9	AB060908	Macaca fa
5	319.2	78.8	4032	9	AK126882	Homo sapi
6	317.6	78.4	4136	9	AK124520	Homo sapi
7	241	59.5	141268	2	AC074365	Homo sapi
8	241	59.5	185467	9	AL606804	Human DNA
9	58	14.3	1532	9	BC030506	Homo sapi
10	58	14.3	1654	9	AF521911	Homo sapi
11	58	14.3	3270	9	AY212246	Homo sapi
12	49	12.1	7218	6	I66494	Sequence 14
13	49	12.1	177229	2	BX855623	Danio fer
14	49	12.1	183117	5	BX004780	Zebrafish
15	48	11.9	4218	9	AK125521	Homo sapi
16	48	11.9	81001	9	AK128688	Homo sapi
17	48	11.9	203405	2	AC024964	Homo sapi
18	48	11.9	214324	2	AC110870	Homo sapi
19	47.2	11.7	1643	6	AX740457	Sequence
20	45.4	11.2	190882	10	AL627105	Mouse DNA
21	44	10.9	1672	10	MMU3263	Mus muscu
22	44	10.9	217409	2	AC130828	Mus muscu
23	43.8	10.8	252980	10	AL669931	Mouse DNA
24	43.4	10.7	69028	9	AC087493	Homo sapi
25	43.4	10.7	124170	9	AC006000	Homo sapi
26	43.4	10.7	186360	9	AC007982	Homo sapi
27	43.4	10.7	218922	4	AC008106	Homo sapi
28	43.2	10.7	372	4	S75319	12S rRNA [n
29	43.2	10.7	176825	10	AC132437	Mus muscu
30	43.2	10.7	199649	10	AL669825	Mouse DNA
31	42.8	10.6	237384	2	AC131424	Rattus no
32	42.8	10.6	244861	2	AC106089	Rattus no
33	42.6	10.5	221012	10	AL672234	Mouse DNA
34	42.6	10.5	268303	2	AL772317	Mus muscu
35	42.6	10.5	294608	2	AC129004	Rattus no
36	42.4	10.5	151316	9	AC005039	Homo sapi
37	42.4	10.5	173924	2	AC102756	Mus muscu
38	42.4	10.5	193629	2	AC139376	Mus muscu
39	42.4	10.5	212480	2	AC135862	Mus muscu
40	42.2	10.4	185872	2	AC145959	Gallus ga
41	42	10.4	171916	2	AC114649	Mus muscu
42	42	10.4	195726	2	AC112925	Mus muscu
43	41.8	10.3	500	5	D15E8A8G1	Dicentrar
44	41.8	10.3	158510	2	AC102478	Mus muscu
45	41.6	10.3	239	11	G37943	C4M29 Plasm

ALIGNMENTS

RESULT 1	BC024174	Homo sapiens	LOC148823	2648 bp	linear	PRI 04-OCT-2003
LOCUS	BC024174	complete cds.				IMAGE:41090641
DEFINITION	BC024174	complete cds.				
ACCESSION	BC024174					
VERSION	BC024174.1	GI:18848218				
KEYWORDS	MGC.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	1 (bases 1 to 2648)					
AUTHORS	Klausner,R.D., Collins,F.S., Wagner,L., Shenmen,C.M., Schuler,G.D.,					

Altschul,S.F., Zeeberg,B., Buetow,K.H., Schaefer,C.F., Bhat,N.K., Hopkins,R.F., Jordan,H., Moore,T., Max,S.I., Wang,J., Hsieh,F., Diatchenko,L., Marusina,K., Farmer,A.A., Rubin,G.M., Hong,L., Stapleton,M., Soares,M.B., Bonaldo,M.F., Casavant,T.L., Schaeetz,T.E., Brownstein,M.J., Usdin,T.B., Toehiyuk,S., Carnici,P., Prange,C., Raha,S.S., Loquellano,N.A., Peters,G.J., Abramson,R.D., Mullany,S.O., Bosak,S.A., McEwan,P.J., McKernan,K.J., Malek,J.A., Gunaratne,P.H., Richards,S., Worley,K.C., Hale,S., Garcia,A.M., Gay,L.J., Hulyk,S.W., Villalon,D.K., Muzny,D.M., Sodergren,E.J., Lu,X., Gibbs,R.A., Fahey,J., Helton,E., Kettelman,M., Madan,A., Rodriguez,S., Sanchez,A., Whiting,M., Madan,A., Young,A.C., Shevchenko,Y., Bouffard,G.G., Blakesley,R.W., Touchman,J.W., Green,E.D., Dickson,M.C., Rodriguez,A.C., Grimwood,J., Schmutz,J., Myers,R.M., Butterfield,S., Krzywinski,M.I., Skalska,U., Smalish,D.E., Schnerch,A., Schein,J.E., Jones,S.J. and Marra,M.A.

Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16999-16903 (2002)

23388257

12477932

2 (Bases 1 to 2648)

Strausberg,R.

Direct Submission

Submitted (19-FEB-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

NIH-MGC Project URL: <http://mgc.nci.nih.gov>

Contact: MGC help desk

Email: cgapbs-r@mail.nih.gov

Tissue Procurement: ATCC

cDNA Library Preparation: CLONTECH Laboratories, Inc.

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)

DNA Sequencing by: Institute for Systems Biology

<http://www.systemsbio.org>

contact: amadansystemsbiology.org

Anup Madan, Jessica Fahey, Erin Helton, Mark Kettelman, Anuradha Madan, Stephanie Rodriguez, Amy Sanchez and Michelle Whiting

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/ILNL at: <http://image.llnl.gov>

Series: IMAGE Plate: 32 Row: b Column: 10

This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 21687148.

FEATURES

source

1..2648

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:24564 IMAGE:4109064"

/tissue_type="Bone marrow, acute myelogenous leukemia"

/clone_lib="NIH MGC_55"

/lab_host="DH10B"

/note="Vector: pDNR-LIB"

1..2648

/gene="LOC148823"

/db_xref="LocusID:148823"

48..455

/codon_start=1

/product="LOC148823"

/protein_id="AAH24174.1"

/db_xref="GI:18848219"

/db_xref="LocusID:148823"

/translation="MGNYLLKLSCLGENQKKPKGNPDEERKQEMTTFERKLQDD KKSQEVSTNQENSGSEEVCTVINIHPQRSLSLSSNDGYNIDSLTRKVRQF RRSSETYALLRTSVSRPCSTHEHDYEVVFPH"

ORIGIN

Query Match 100.0%; Score 405; DB 9; Length 2648;

Best Local Similarity 100.0%; Pred. No. 9.5e-94;

Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 60

DB 48 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 107

QY 61 AAAGGAAATCCAGATGAGAAAGAAACCGGAGGAATGACTACATTTGAAAGAAACTT 120

DB 108 AAAGGAAATCCAGATGAGAAAGAAACCGGAGGAATGACTACATTTGAAAGAAACTT 167

QY 121 CAAGATCAAGATAAGAAAGCAAGAACTTTCATCCACTTCTAATCAGGAAACGAGAAT 180

DB 168 CAAGATCAAGATAAGAAAGCAAGAACTTTCATCCACTTCTAATCAGGAAACGAGAAT 227

QY 181 GCACATGGTTCTGAAGAGTGTCTACTGTCTATTAATCATCATCCCCCATCAGAGATCC 240

DB 228 GCACATGGTTCTGAAGAGTGTCTACTGTCTATTAATCATCATCCCCCATCAGAGATCC 287

QY 241 TCCTCTGAGCTCCAAATGATGGCTATCAGAAACATTTGACTCCCTCACAGGAAAGTGAGA 300

DB 288 TCCTCTGAGCTCCAAATGATGGCTATCAGAAACATTTGACTCCCTCACAGGAAAGTGAGA 347

QY 301 CAGTTTAGAGAAAGTTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGAGGCT 360

DB 348 CAGTTTAGAGAAAGTTCAGAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGAGGCT 407

QY 361 TCTTCTCTGACCCATCAGCATGATTATCAAGTTGTGTTTCCACAC 405

DB 408 TCTTCTCTGACCCATCAGCATGATTATCAAGTTGTGTTTCCACAC 452

RESULT 2

AX505122

LOCUS AX505122 3762 bp DNA linear PAT 27-SEP-2002

DEFINITION Sequence 1 from Patent WO0246389.

ACCESSION AX505122

VERSION AX505122.1 GI:23386429

KEYWORDS Homo sapiens (human)

SOURCE Homo sapiens

ORGANISM Homo sapiens

REFERENCE 1

AUTHORS Nocka,K., Pirozzi,G. and Einstein,R.

TITLE Novel genes associated with allergic hypersensitivity and mast cell activation

JOURNAL Patent: WO 0246389-A 1 13-JUN-2002;

UCB, S.A. (BE)

FEATURES

Location/Qualifiers

1..3762

/organism="Homo sapiens"

/mol_type="unassigned DNA"

/db_xref="taxon:9606"

25..432

/note="unnamed protein product"

/codon_start=1

/protein_id="CAD48808.1"

/db_xref="GI:23386430"

/db_xref="REMBL:CAD48808"

/translation="MGNYLLKLSCLGENQKKPKGNPDEERKQEMTTFERKLQDD KKSQEVSTNQENSGSEEVCTVINIHPQRSLSLSSNDGYNIDSLTRKVRQF RRSSETYALLRTSVSRPCSTHEHDYEVVFPH"

ORIGIN

Query Match 100.0%; Score 405; DB 6; Length 3762;

Best Local Similarity 100.0%; Pred. No. 9.1e-94;

Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 60

DB 25 ATGGGAAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 84

QY 61 AAAGGAAATCCAGATGAGAAAGAAACCGGAGGAATGACTACATTTGAAAGAAACTT 120

DB 85 AAAGGAAATCCAGATGAGAAAGAAACCGGAGGAATGACTACATTTGAAAGAAACTT 144

```

QY 121 CAAGATCAAGTAAAGAACCAAGAGTTTTCATCCACTTCTTAATCAGGAAAACGAGAAAT 180
Db 145 CAAGATCAAGTAAAGAACCAAGAGTTTTCATCCACTTCTTAATCAGGAAAACGAGAAAT 204
QY 181 GCGAGTGGTCTGAGAAAGTGTGTACACTGTCTAATTAATCAGATCCCCCATCAGAGATCC 240
Db 205 GCGAGTGGTCTGAGAAAGTGTGTACACTGTCTAATTAATCAGATCCCCCATCAGAGATCC 264
QY 241 TCCCTGAGCTCCATGATGATGGCTATGAGAACATTCGACTCCCTCAAGAGAAAGTGAGA 300
Db 265 TCCCTGAGCTCCATGATGATGGCTATGAGAACATTCGACTCCCTCAAGAGAAAGTGAGA 324
QY 301 CAGTTTAGAGAAAGGTCAGACAGAGAATATGCCCTTCTTAGGACTTCTGTAGTAGGCCT 360
Db 325 CAGTTTAGAGAAAGGTCAGACAGAGAATATGCCCTTCTTAGGACTTCTGTAGTAGGCCT 384
QY 361 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCCACAC 405
Db 385 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCCACAC 429

RESULT 3
AKI23798 1709 bp mRNA linear PRI 09-SEP-2003
LOCUS Homo sapiens cDNA FLJ1804 fis, clone NOV2000710.
DEFINITION AKI23798
VERSION AKI23798.1 GI:34529425
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1

Suzuki, O., Sasaki, N., Aotsuka, S., Shoji, T., Ichihara, T.,
Shiohata, N., Matsumoto, K., Hirano, M., Sano, S., Nomura, R.,
Yoshihikawa, Y., Matsumura, Y., Moriya, S., Chiba, E., Momiyama, H.,
Oogawa, S., Kaeriyama, S., Sato, N., Matsunawa, H., Takahashi, E.,
Kataoka, R., Kuga, N., Kuroda, A., Sato, I., Kanata, K., Takami, S.,
Terashima, Y., Watanabe, M., Sugiyama, T., Irie, R., Otsuki, T.,
Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Ohshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 1709)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI) (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
Location/Qualifiers
1. .1709
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NOV2000710"
/tissue_type="ovary"
/clone_lib="NOVAR2"
/dev_stage="adult"
/notes="cloning vector: pME18SFL3"

FEATURES
source
1. .1709
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="NOV2000710"
/tissue_type="ovary"
/clone_lib="NOVAR2"
/dev_stage="adult"
/notes="cloning vector: pME18SFL3"

ORIGIN

```

```

Query Match 99.6%; Score 403.4; DB 9; Length 1709;
Best Local Similarity 99.8%; Pred. No. 2.6e-93;
Matches 404; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATGGGAATATATCTCTCGGAAAACCTCAGTTGCTGGGAGAGAAATCAAAAAGGCCCAAG 60
Db 59 ATGGGAATATATCTCTCGGAAAACCTCAGTTGCTGGGAGAGAAATCAAAAAGGCCCAAG 118
QY 61 AAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAAACCTT 120
Db 119 AAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAAACCTT 178
QY 121 CAAGATCAAGATAAGAAAAGCCAAAGACTTTTCATCCACTTCTTAATCAGGAAAACGAGAAAT 180
Db 179 CAAGATCAAGATAAGAAAAGCCAAAGACTTTTCATCCACTTCTTAATCAGGAAAACGAGAAAT 238
QY 181 GGCAGTGGTCTGAGAGAAAGTGTGTACACTGTCTAATTAATCAGATCCCCCATCAGAGATCC 240
Db 239 GGCAGTGGTCTGAGAGAAAGTGTGTACACTGTCTAATTAATCAGATCCCCCATCAGAGATCC 298
QY 241 TCCCTGAGCTCCATGATGATGGCTATGAGAACATTCGACTCCCTCAAGAGAAAGTGAGA 300
Db 299 TCCCTGAGCTCCATGATGATGGCTATGAGAACATTCGACTCCCTCAAGAGAAAGTGAGA 358
QY 301 CAGTTTAGAGAAAGGTCAGACAGAGAATATGCCCTTCTTAGGACTTCTGTAGTAGGCCT 360
Db 359 CAGTTTAGAGAAAGGTCAGACAGAGAATATGCCCTTCTTAGGACTTCTGTAGTAGGCCT 418
QY 361 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCCACAC 405
Db 419 TGTTCCTGCACCCATGAGCATGATTATGAAGTTGTGTTCCACAC 463

RESULT 4
AB060908 1924 bp mRNA linear PRI 10-JAN-2002
LOCUS Macaca fascicularis brain cDNA clone: Qtra-14007, full insert
DEFINITION AB060908
ACCESSION AB060908
VERSION AB060908.1 GI:13874585
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Macaca fascicularis (crab-eating macaque)
ORGANISM Macaca fascicularis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Cercopitheidae;
Cercopitheciinae; Macaca.
1

Osada, N., Hida, M., Kusuda, J., Tanuma, R., Hirata, M., Hirai, M.,
Terao, K., Suzuki, Y., Sugano, S. and Hashimoto, K.
Prediction of unidentified human genes on the basis of sequence
similarity to novel cDNAs from cynomolgus monkey brain
Online Publication
Genome Biology 2001 3 (1): research0006.1-0006.5;
http://genomebiology.com/2001/3/1/research/0006/
2 (bases 1 to 1924)
Hashimoto, K., Osada, N., Hida, M., Kusuda, J. and Sugano, S.
Direct Submission
Submitted (27-APR-2001) Katsuyuki Hashimoto, National Institute of
Infectious Diseases, Division of Genetic Resources; 23-1, Toyama
1-chome, Shinjuku-ku, Tokyo 162-8640, Japan
(E-mail: khashi@nih.go.jp, URL: http://www.nih.go.jp/yoken/genebank/,
Tel: 81-3-5285-1111 (ex. 2120), Fax: 81-3-5285-1181)
Lab host: TOP10
Vector: pME18S-FL3 (Acc. No. AB009864)
R. Site1: DraIII (CACTGTGTG)
R. Site2: DraIII (CACCATGTG)
Description: 1st strand cDNA was primed with an oligo(dT) primer
[ATGCGCCCTTTTCTTTTCTTTT]; double-stranded cDNA was synthesized
using specific 5' and 3' primers and amplified by PCR. The PCR
product was digested with SfiI and size selection was performed to
exclude fragments <1.5kb. The SfiI-digested PCR product was cloned
into distinct DraIII sites of pME18S-FL3. XhoI sites just outside

```

the DralII sites can be used to isolate the cDNA insert. Libraries were constructed by oligo-capping method (Sugano et al., Institute of Medical Science, University of Tokyo).

Custom primer used for sequencing
(5' end primer [CTCTGCTCAAGACTCGG];
3' end primer [CGACCTGCTCAGCACA]).

FEATURES

source

```

1.1924
/organism="Macaca fascicularis"
/mol_type="mRNA"
/db_xref="taxon:9541"
/clone="Qcra-14007"
/sex="male"
/tissue="temporal lobe right"
/tissue_type="temporal lobe right"
/clone_lib="macaque brain cDNA library Qcra"
/clone_stage="adult"
248..559
/codon_start=1
/product="hypothetical protein"
/protein_id="BA46903.1"
/db_xref="GI:13874586"
/translations="MTTFERKLDQDKSKEVSSISQENENGSGSEVCYTVINHIP
HRRSSLSNDGVENIDSLTRKVRERSETEYALLRTSVSRPYSCTHEHDYEVVLP
H"

```

ORIGIN

```

Query Match      88.7%; Score 359.4; DB 9; Length 1924;
Best Local Similarity 94.7%; Pred. No. 5.5e-82;
Matches 372; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 13 CTCCTCGGAAACCTCAGTTGCTGGGAGAGAAATCAAAAGAACGCCAAGAAAGAAACCCCA 72
Db 164 CGCGGGAGGACATGAGTTGCTTGGGAGAGAAATCAAAAGAACGCCAAGAAAGAAACCCCA 223

QY 73 GATGAGGAAGAAACCGCGAGGAATGACTACATTTGAAGAAATCTCAAGATCAAGAT 132
Db 224 GATGAGGAAGAAACCGCGAGGAATGACTACATTTGAAGAAATCTCAAGATCAAGAT 283

QY 133 AAGAAAGCCAGAAAGTTTCAATCAGGAAACGAGAAATGCGAGTGGTGTCT 192
Db 284 AAGAAAGCCAGAAAGTTTCAATCAGGAAACGAGAAATGCGAGTGGTGTCT 343

QY 193 GAAGAAGTGCTACATGCTATTAATCAATCCCATCAGAGATCCCTCCCTGAGCTCC 252
Db 344 GAAGAAGTGCTACATGCTATTAATCAATCCCATCAGAGATCCCTCCCTGAGCTCC 403

QY 253 AATGATGCTGATGAGAAATGCTCCCTCAGAGAAAGTGAGACAGTTTAGAGAA 312
Db 404 AATGATGCTGATGAGAAATGCTCCCTCAGAGAAAGTGAGACAGTTTAGAGAA 463

QY 313 AGGTCAGAGACAGAAATATGCTTTTGAAGATCTTGTAGAGGCTTTGTTCTGCACC 372
Db 464 AGGTCAGAGACAGAAATATGCTTTTGAAGATCTTGTAGAGGCTTTGTTCTGCACC 523

QY 373 CATGAGCATGATATGAGTTGTTTCCACAC 405
Db 524 CATGAGCATGATATGAGTTGTTTCCACAC 556

```

RESULT 5

```

AKI26682      4032 bp mRNA linear PRI 09-SEP-2003
LOCUS
DEFINITION Homo sapiens cDNA FLJ44728 fis, clone BRACE3024537.
ACCESSION AKI26682
VERSION AKI26682.1 GI:34533254
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

REFERENCE

1. Ninomiya, K., Wagatsuma, M., Kanda, K., Kondo, H., Yokoi, T.,

Kodaira, H., Furuya, T., Takahashi, M., Kikkawa, E., Omura, Y., Abe, K.,
Kanihara, K., Katsuka, N., Sato, K., Tanikawa, M., Yamazaki, M.,
Sugiyama, T., Irie, R., Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S.,
Yamamoto, J., Isono, Y., Kawai-Hio, Y., Saito, K., Nishikawa, T.,
Kimura, K., Yamashita, H., Matsuo, K., Nakamura, Y., Sekine, M.,
Kikuchi, H., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahari, K., Masuho, Y., Nagai, K. and Isogai, T.
NEO human cDNA sequencing project

TITLE

JOURNAL

REFERENCE

AUTHORS

JOURNAL

COMMENT

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

Location/Qualifiers

source

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

Location/Qualifiers

source

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

Location/Qualifiers

source

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

Location/Qualifiers

source

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

Location/Qualifiers

source

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

Location/Qualifiers

source

1. 4032

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="BRACE3024537"

/tissue_type="cerebellum"

/clone_lib="BRACE3"

/note="Cloning vector: pME18SFL3"

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1
REFERENCE
AUTHORS
TITLE
JOURNAL

2
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL

COMMENT

FEATURES
source
1. 4136
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="BRACE3002390"
/tissue_type="cerebellum"
/clone_lib="BRACE3"
/note="cloning vector: pME18SFL3"
136. 537
/note="unnamed protein product"
/codon_start=1
/protein_id="BAC85873.1"
/db_xref="GI:34530325"
/translation="MWMLEDSNSTGCGARNCVCLVFLSVLGCQSRKQMTQQAG
RWLRAGREASSETPEGLTLDQFSLPAAGWQHPVRCNVLPHPAFSSWCFQG
NFPKWSQELNRVTVKRWELIISGSRK"

ORIGIN
Query Match 78.4%; Score 317.6; DB 9; Length 4136;
Best Local Similarity 94.5%; Pred. No. 2.9e-71;
Matches 329; Conservative 0; Mismatches 19; Indels 0; Gaps 0;
QY 58 AAGAAAGGAAACCCAGATGAGGAAGAAACCGCAGGAATGACTACATTGTAAGAAAA 117
DB 495 AAGATGGGAATATTCTCTCGAAACTCAGGCAGGAATGACTACATTGTAAGAAAA 554
QY 118 CTTCAAGATCAAGTAAGAAAGCCAGAGTTTCATCCACTTCTATCAGGAACGAG 177
DB 555 CTTCAAGATCAAGTAAGAAAGCCAGAGTTTCATCCACTTCTATCAGGAACGAG 614
QY 178 AATGCACTGGTCTTGAAGAAGTGTGCTACACTGTCTATTAATCATCCCCCATCAGAGA 237
DB 615 AATGCACTGGTCTTGAAGAAGTGTGCTACACTGTCTATTAATCATCCCCCATCAGAGA 674
QY 238 TCCTCCCTGAGTCCAGTATGATGGCTATGAGACATGACTGCTCCCTCAGAGAAAGTG 297
DB 675 TCCTCCCTGAGTCCAGTATGATGGCTATGAGACATGACTGCTCCCTCAGAGAAAGTG 734
QY 298 AGACAGTTTAGAAGAGTCCAGACAGAAATATGCCCTTCTTAGCACTTCTGTAGTAGG 357
DB 735 AGACAGTTTAGAAGAGTCCAGACAGAAATATGCCCTTCTTAGCACTTCTGTAGTAGG 794
QY 358 CTTGTTCTCCTGACCCAGCATGATTAATGAAGTTGTGTTCCACAC 405

Db 795 CCTGTGCTCTGCACCATGAGCATGATTATGAAGTTGTGTTCCACAC 842

RESULT 7
AC074365/c
LOCUS
DEFINITION
AC074365
AC074365.5 GI:10280935
HTG; HTGS_PHASE1; HTGS_DRAFT.
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

AC074365
Homo sapiens chromosome 1 clone RP11-115C4, WORKING DRAFT SEQUENCE,
10 unordered pieces.
AC074365
HTG; HTGS_PHASE1; HTGS_DRAFT.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Waterston,R.H.
Unpublished
2 (bases 1 to 141268)
Waterston,R.H.
Direct Submission
Submitted (29-JUL-2000) Genome Sequencing Center, Washington
University School of Medicine, 444 Forest Park Parkway, St. Louis,
MO 63108, USA
On Sep 23, 2000 this sequence version replaced gi:9838075.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H.NH0115C04
----- Summary Statistics -----
Sequencing vector: pMI3; 100%
Sequencing method: Dye-terminator Big Dye; 0% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 136274 bases at least Q40
Consensus quality: 138356 bases at least Q30
Consensus quality: 139182 bases at least Q20
Insert size: 142000; agarose-fp
Insert size: 140368; sum-of-contigs
Quality coverage: 4.64 in Q20 bases; agarose-fp
Quality coverage: 4.81 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 10 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
* 1 31377: contig of 31377 bp in length
* 31378 31477: gap of unknown length
* 31478 63011: contig of 31534 bp in length
* 63011 63111: gap of unknown length
* 63111 64845: contig of 1734 bp in length
* 64845 64945: gap of unknown length
* 64945 67606: contig of 2661 bp in length
* 67606 67707: gap of unknown length
* 67707 77674: contig of 9968 bp in length
* 77674 89256: contig of 11482 bp in length
* 89256 89357: gap of unknown length
* 89357 100401: contig of 11044 bp in length
* 100401 100501: gap of unknown length
* 100501 111915: contig of 11415 bp in length
* 111915 112015: gap of unknown length
* 112015 125713: contig of 13698 bp in length
* 125713 125813: gap of unknown length

* 125814 141268: contig of 15455 bp in length.

```

FEATURES
  source
    Location/Qualifiers
      1..141268
        /organism="Homo sapiens"
        /mol_type="genomic DNA"
        /db_xref="taxon:9606"
        /chromosome="1"
        /clone="RP11-115C4"
      1..31377
        /note="assembly_name:Contig10"
      31478..63011
        /note="assembly_name:Contig11"
      63112..64845
        /note="assembly_name:Contig2"
      64946..67606
        /note="assembly_name:Contig3"
      67707..77674
        /note="assembly_name:Contig4"
      77775..89256
        /note="assembly_name:Contig5"
      89357..100400
        /note="assembly_name:Contig6"
      100501..111915
        /note="assembly_name:Contig7"
      112016..125713
        /note="assembly_name:Contig8"
      125814..141268
        /note="assembly_name:Contig9"
        clone_end:SP6
        vector_side:left"

ORIGIN
Query Match      59.5%; Score 241; DB 2; Length 141268;
Best Local Similarity 96.1%; Pred. No. 9.4e-52;
Matches 247; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 149 TTTCATCCACTTCTAATCAGGAAACGAGATGGCAGTGGTCTTGAAGAAGTGTGCTACA 208
Db 5717 TTTCCTTTGTGCTTGGCAGGAAACGAGATGGCAGTGGTCTTGAAGAAGTGTGCTACA 5658

QY 209 CTGTCATTAAACACATCCCCCATCAGAGATCTCCCTGAGTCCCAATGATGCGCTATG 269
Db 5657 CTGTCATTAAACACATCCCCCATCAGAGATCTCCCTGAGTCCCAATGATGCGCTATG 5598

QY 269 AGAACATTGACTCCCTCAACAGGAAGTGAGACATCTTTAGAGAAGGTTCAGACAGAGAT 328
Db 5597 AGAACATTGACTCCCTCAACAGGAAGTGAGACATCTTTAGAGAAGGTTCAGACAGAGAT 5538

QY 329 ATGCCCTTCTTAGGACTTCTGTAGTGGCTTGTCTGACCCCATGAGCATGATTATG 388
Db 5537 ATGCCCTTCTTAGGACTTCTGTAGTGGCTTGTCTGACCCCATGAGCATGATTATG 5478

QY 389 AAGTTGTGTTTCCACAC 405
Db 5477 AAGTTGTGTTTCCACAC 5461

RESULT 8
AL606804 185467 bp DNA linear PRI 23-OCT-2002
LOCUS Human DNA sequence from clone RP11-978115 on chromosome 1, complete
DEFINITION
ACCESSION AL606804
VERSION AL606804.11 GI:24366459
KEYWORDS HTG.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominiidae; Homo.
REFERENCE 1 (bases 1 to 185467)
AUTHORS Almeida, J.
TITLE Direct Submission
JOURNAL Submitted (16-OCT-2002) Wellcome Trust Sanger Institute, Hinxton,

```

COMMENT

Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
 On Oct 24, 2002 this sequence version replaced gi:18121511.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; Sw, SWISSPROT; Tr, TREMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at

http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 1, constructed by the Sanger Centre Chromosome 1 Mapping Group. Further information can be found at

http://www.sanger.ac.uk/HGP/Chr1
 RP11-978115 is from the library RP11-11.4 constructed by the group of Fietier de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pBAC3.6

----- Genome Center

Center: Wellcome Trust Sanger Institute

Center code: SC

Web site: http://www.sanger.ac.uk

Contact: humquery@sanger.ac.uk

FEATURES

Location/Qualifiers

source

1..185467

/organism="Homo sapiens"

/mol_type="genomic DNA"

/db_xref="taxon:9606"

/chromosome="1"

/clone="RP11-978115"

/clone_lib="RP11-11.4"

Query Match 59.5%; Score 241; DB 9; Length 185467;

Best Local Similarity 96.1%; Pred. No. 9e-52;

Matches 247; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

QY 149 TTTCATCCACTTCTAATCAGGAAACGAGATGGCAGTGGTCTTGAAGAAGTGTGCTACA 208

Db 132063 TTTCCTTTGTGCTTGGCAGGAAACGAGATGGCAGTGGTCTTGAAGAAGTGTGCTACA 132122

QY 209 CTGTCATTAAACACATCCCCCATCAGAGATCTCCCTGAGTCCCAATGATGCGCTATG 268

Db 132123 CTGTCATTAAACACATCCCCCATCAGAGATCTCCCTGAGTCCCAATGATGCGCTATG 132182

QY 269 AGAACATTGACTCCCTCAACAGGAAGTGAGACATCTTTAGAGAAGGTTCAGACAGAGAT 328

Db 132183 AGAACATTGACTCCCTCAACAGGAAGTGAGACATCTTTAGAGAAGGTTCAGACAGAGAT 132242

QY 329 ATGCCCTTCTTAGGACTTCTGTAGTGGCTTGTCTGACCCCATGAGCATGATTATG 388

Db 132243 ATGCCCTTCTTAGGACTTCTGTAGTGGCTTGTCTGACCCCATGAGCATGATTATG 132302

QY 389 AAGTTGTGTTTCCACAC 405

Db 132303 AAGTTGTGTTTCCACAC 132319

RESULT 9

BC030506

LOCUS

BC030506

1532 bp mRNA linear PRI 06-OCT-2003

DEFINITION Homo sapiens germinal center expressed transcript 2, mRNA (CDNA clone MGC:40441 IMAGE:4385178), complete cds.

ACCESSION BC030506

VERSION 1

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 1532)

AUTHORS Strausberg, R.L., Feigold, E.A., Grouse, L.H., Derge, J.G., Klauener, R.D., Collins, F.S., Wagner, L., Schmen, C.M., Schuler, G.D., Altschul, S.F., Zebberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K., Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F., Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L., Stapleton, M., Soares, M.B., Bonaldo, M.P., Casavant, T.L., Scheetz, T.E., Brownstein, M.J., Ussid, T.B., Toshiyuki, S., Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J., Abramson, R.D., Mullany, S.J., Bosak, S.A., McSwan, P.J., McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S., Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hulyk, S.W., Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A., Rahe, J., Helton, E., Kettelman, M., Madan, A., Rodriguez, S., Sanchez, A., Whiting, M., Madan, A., Young, A.C., Shevchenko, Y., Bouffard, G.G., Blackesley, R.W., Touchman, J.W., Green, E.D., Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M., Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smal, D.E., Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.

TITLE Generation and initial analysis of more than 15,000 full-length human and mouse cDNA sequences

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)

MEDLINE 2238257

PUBMED 12477932

REFERENCE 2 (bases 1 to 1532)

AUTHORS Strausberg, R.

TITLE Direct Submission

JOURNAL Submitted (07-MAY-2002) National Institutes of Health, Mammalian Gene Collection (MGC), Cancer Genomics Office, National Cancer Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590, USA

REMARK NIH-MGC Project URL: <http://mgc.nci.nih.gov>

COMMENT Contact: MGC help desk Email: cgabs-remail.nih.gov Tissue Procurement: Lou Staudt cDNA Library Preparation: Life Technologies, Inc. cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL) DNA Sequencing by: National Institutes of Health Intramural Sequencing Center (NISC), Gaithersburg, Maryland Web site: <http://www.nisc.nih.gov/> Contact: nisc_mgc@hgr.nih.gov Akhter, N., Ayale, K., Beckstrom-Sternberg, S.M., Benjamin, B., Blakesley, R.W., Bouffard, G.G., Breen, K., Brinkley, C., Brooks, S., Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Heghigni, P., Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R., Maduro, Q.L., Masello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C., McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W., Tsurgon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L., Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov> Series: IRAC Plate: 64 Row: a Column: 15 This clone was selected for full length sequencing because it passed the following selection criteria: matched mRNA gi: 22749536. Location/Qualifiers

1..1532

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="MGC:40441 IMAGE:4385178"

/tissue_type="Lymph, lymphoma"

/clone_lib="NTH_MGC_85"

gene

lab_host="DH10B"

/note="Vector: pCMV-SPORT6"

1..1532

/gene="GCET2"

/note="synonyms: HGAL, GCAT2, MGC40441"

/db_xref="locusid:257144"

/db_xref="MIM:607792"

54..590

/codon_start=1

/product="germinal center expressed transcript 2"

/protein_id="AAH30506.1"

/db_xref="GI:20987806"

/db_xref="LocusID:257144"

/translation="MGSLRLKRNKRNQNTQEMFWNQSPKORTSRCHWHDIAEGCF CLPKKILIFKQDSQNRMSSTPIQNDVDTSEELCYLLINHRVICTIPSGMS AEEYENVPCKAERPSLGGTETETYSLLHMPSTDPHAKSPDEYELLMPLRISHP LQQRPLMAPSETQFSLH"

CDS

ORIGIN

Query Match 14.3%; Score 58; DB 9; Length 1532;

Best Local Similarity 54.4%; Pred. No. 0.00023;

Matches 162; Conservative 0; Mismatches 130; Indels 6; Gaps 2;

QY 114 AAAACTTCAAGATCAAGATAGAAAGCAAGAAAGTTTCATCCACTTCTATCATCAGGAAA 173

Db 218 AAAGAGGCAAGATTCCTCCAAACGAAATGAAGATGTCACTACTCCATCCAGGACAA 277

QY 174 CGAGATGGCAGTGGTCTTGAAGAAGTGTCTACACTGTCTAATATCACA---TCCCCCA 230

Db 278 TGTTCACAGACCTACTCAGAGGAGCTGTGTATACCTCATCAATATCGGGTCTCTG 337

QY 231 TCAGAGATCTCCCTGAGCTCCATGATGATGGTGTATGAGAACTTGAATCTCCCTCACAAG 290

Db 338 TACAAGGCAATCAGGGAATCTGCTGAAGACTATGAGAAATGTCCTGCAAGCTGA 397

QY 291 GAAATGAGACAGTTTGAAGAAAGTTCAGACAGATATGCCCTTCT---TAGACTTC 347

Db 338 GAGACCCAGAGAGTCTTTGGGAGGAACCTGAGACTGAGTATTCATCTTACATATGCTTC 457

QY 348 TGTTAGTAGGCTTCTCTCTGACCCATGAGTATGATGAAGTTGTGTTTCCACAC 405

Db 458 TACAGACCCAGGCAATGCCATCCCAAGAGATGATGACTTCTCATGCTCTCAC 515

RESULT 10

AF521911

LOCUS

DEFINITION Homo sapiens HGAL mRNA, complete cds.

ACCESSION AF521911

VERSION AF521911.1

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1654)

Loscos, I.S., Allazadeh, A., Rajapaksa, R., Tibshirani, R. and Levy, R.

AUTHORS HGAL is a novel interleukin-4-inducible gene that strongly predicts survival in diffuse large B-cell lymphoma

TITLE

JOURNAL Blood 101 (2), 433-440 (2003)

MEDLINE 22397484

PUBMED 12509382

REFERENCE 2 (bases 1 to 1654)

Loscos, I.S., Allazadeh, A. and Levy, R.

AUTHORS Direct Submission

TITLE Submitted (16-JUN-2002) Division of Oncology, Department of Medicine, Stanford University, 269 Campus Drive, CCSR building, Room 1100, Palo Alto, CA 94305-5151, USA

JOURNAL

Location/Qualifiers

1..1654

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

misc_feature /note="assembly fragment:00920
fragment chain:2"
97311..103462
misc_feature /note="assembly fragment:00210
fragment chain:2"
103563..107200
misc_feature /note="assembly fragment:00054
fragment chain:2"
107301..111003
misc_feature /note="assembly fragment:00117
fragment chain:2"
11104..128252
misc_feature /note="assembly fragment:01194
fragment chain:2"
128353..134350
misc_feature /note="assembly fragment:00307
fragment chain:2"
134351..141346
misc_feature /note="assembly fragment:00340
fragment chain:2"
141447..146748
misc_feature /note="assembly fragment:00242
fragment chain:2"
146849..149324
misc_feature /note="assembly fragment:00023
fragment chain:2"
149425..152942
misc_feature /note="assembly fragment:00098
fragment chain:2"
153043..164529
misc_feature /note="assembly fragment:00529
fragment chain:2"
164630..172404
misc_feature /note="assembly fragment:00392
fragment chain:3"
172505..174733
misc_feature /note="assembly fragment:00082
fragment chain:3"
174834..177229
misc_feature /note="assembly fragment:00044"

ORIGIN

Query Match 12.1%; Score 49; DB 2; Length 177229;

Best Local Similarity 59.9%; Pred. No. 0.025;

Matches 82; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1 ATGGGAATTTATCTCTCGGAAATCTAGTTGCCTGGGAGAGATCAAAAGAGCCCAAG 60
DB 144111 AAGACAATAACCACTCGATACACTGAGTTGTGCGAAGAGCATCTCTAAAGCACAAC 144170
QY 61 AAAGGAACCCAGATCAGGAAGAAACCGGAGGAATGACTATACATTTGAAAGAAACTT 120
DB 144171 ACATGAACCTTGAGCGGATGGCTACAGCGCAAGACCACTCTGTCGGCTAAGAA 144230
QY 121 CAAGATCAAGATAAGAA 137
DB 144231 CAGGAACAGGCAAGAA 144247

RESULT 14

EX004780/c 188317 bp DNA linear VRT 13-MAR-2003
LOCUS Zebrafish DNA sequence from clone CH211-202P1 in linkage group 20,
DEFINITION complete sequence.
ACCESSION BX004780
VERSION BX004780.4 GI:28552191
KEYWORDS HTG
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE 1 (bases 1 to 188317)

AUTHORS

Direct Submission
Submitted (13-MAR-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
zfsh-help@sanger.ac.uk Clone requests: clonerequest@sanger.ac.uk

COMMENT

On Feb 24, 2003 this sequence version replaced gi:28460278.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: zfsh-help@sanger.ac.uk

During sequence assembly data is compared from overlapping clones.
Where differences are found these are annotated as variations
together with a note of the overlapping clone name. Note that the
variation annotation may not be found in the sequence submission
corresponding to the overlapping clone, as we submit sequences with
only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all
regions were either double-stranded or sequenced with an alternate
chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such
as compressions and repeats; all regions were covered by at least
one plasmid subclone or more than one M13 subclone; and the
assembly was confirmed by restriction digest, except on the rare
occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession
numbers given in the feature table with their source databases:
Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information
on the WORMPEP database can be found at
http://www.sanger.ac.uk/projects/C.elegans/wormpep Repeat names
beginning 'Dr' were identified by The Recon repeat discovery system
(Zhirong Bao and Sean Eddy, submitted), and those beginning 'drr'
were identified by Rick Waterman (Stephen Johnson lab, WashU). For
further information see http://www/projects/d_rerio/fishmask.shtml
CH211-202P1 is from a CHORI-211 BAC library
VECTOR: pTARBAC2.1.

FEATURES

Location/Qualifiers
1..188317
/organism="Danio rerio"
/mol_type="genomic DNA"
/db_xref="taxon:7955"
/clone="CH211-202P1"
/clone_lib="CHORI-211"

ORIGIN

Query Match 12.1%; Score 49; DB 5; Length 188317;
Best Local Similarity 59.9%; Pred. No. 0.025;
Matches 82; Conservative 0; Mismatches 55; Indels 0; Gaps 0;
QY 1 ATGGGAATTTATCTCTCGGAAATCTAGTTGCCTGGGAGAGATCAAAAGAGCCCAAG 60
DB 120337 AAGACAATAACCACTCGATACACTGAGTTGTGCGAAGAGCATCTCTAAAGCACAAC 120178
QY 61 AAAGGAACCCAGATCAGGAAGAAACCGGAGGAATGACTATACATTTGAAAGAAACTT 120
DB 120177 ACATGAACCTTGAGCGGATGGCTACAGCGCAAGACCACTCTGTCGGCTAAGAA 120118
QY 121 CAAGATCAAGATAAGAA 137
DB 120117 CAGGAACAGGCAAGAA 120101

RESULT 15

AK125521/c 4218 bp mRNA linear PRI 09-SEP-2003
LOCUS Homo sapiens CDNA FLJ43533 fis, clone PLACE7001544.
DEFINITION AK125521
ACCESSION AK125521
VERSION AK125521.1 GI:34531645
KEYWORDS oligo capping; fis (full insert sequence).
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 Tashiro, H., Yamazaki, M., Watanabe, K., Kumaqai, A., Itakura, S.,
Fukuzumi, Y., Fujimori, Y., Komiyama, M., Sugiyama, T., Irie, R.,
Otsuki, T., Sato, H., Wakamatsu, A., Ishii, S., Yamamoto, J., Isono, Y.,
Kawai-Hio, Y., Saito, K., Nishikawa, T., Kimura, K., Yamashita, H.,
Matsuo, K., Nakamura, Y., Sekine, M., Kikuchi, H., Kanda, K.,
Wagatsuma, M., Murakawa, K., Kanehori, K., Takahashi-Fujii, A.,
Oshima, A., Sugiyama, A., Kawakami, B., Suzuki, Y., Sugano, S.,
Nagahara, K., Masuho, Y., Nagai, K. and Isogai, T.
NEDO human cDNA sequencing project
Unpublished
2 (bases 1 to 4218)
Isogai, T. and Yamamoto, J.
Direct Submission
Submitted (15-JUL-2003) Takao Isogai, FLJ Project (HRI Team); 2-6-7
Kazusa-Kamatari, Kisarazu, Chiba 292-0818, Japan
(E-mail: genomics@hri.co.jp, Tel: 81-438-52-3975, Fax: 81-438-52-3986)
NEDO human cDNA sequencing project supported by Ministry of
Economy, Trade and Industry of Japan; cDNA full insert sequencing:
Research Association for Biotechnology (RAB); cDNA library
construction: Helix Research Institute (HRI); (supported by Japan
Key Technology Center etc.); 5'- & 3'-end one pass sequencing: RAB,
HRI, and Biotechnology Center, National Institute of Technology and
Evaluation; clone selection for full insert sequencing: HRI and
RAB; annotation: HRI and RAB.
FEATURES
Location/Qualifiers
1..4218
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="PLACE7001544"
/tissue_type="Placenta"
/clone_lib="PLACE7"
/note="cloning vector: pME18SFL3"

ORIGIN

Query Match 11.9%; Score 48; DB 9; Length 4218;
Best Local Similarity 52.0%; Pred. No. 0.075;
Matches 158; Conservative 0; Mismatches 140; Indels 6; Gaps 2;

QY 108 TGAAGAAACCTTCAAGATCAAGATAGAAAGCCAGAGTTTCATCCACTTCTATCA 167
DB 3442 TGAATGTAAGGTCATGGCAAGAGACCAATCTTGATGGTCTTCTTTGACAAGCA 3383
QY 168 GGAAACGAGAGATGCGAGTGTCTGAAGAAGTGCTACACTGTCAATATCA--CAT 224
DB 3382 GGACATGTTGACAGAGCTTACTAGAGAGCTGTGCTATACCTCATCAATCATCGGGT 3323
QY 225 CCCCATCAGAGATCCTCCCTGAGTCCCAATGATGANGGCTATGAGAACATGACTCCCT 284
DB 3322 TCTCTGTAAGGCCATCAGGGAAGTCTGCTGAGAGTACTATGAGAATGTTCCCTGCAA 3263
QY 285 CACAAGGAAGTGAGACAGTTTAGAAGAGGTGAGAGACAGATATGCCCTTCTAGGAC 344
DB 3262 AGCTGAGAGACCCAGAGAGTCCCTGGGAGGAAGTACTGAGTATTCACCTTCTACATAT 3203
QY 345 TTCTGTTAGTGGCCCTTG---TTCTGCAACCCATGAGCATGATTATGAAGTTGTGTTCC 401
DB 3202 GCCTTCTACAGACCCAGGCGATGCCGATGCCAGAGATGAATGAACTTCTCATGCC 3143
QY 402 ACAC 405
DB 3142 TCAC 3139

Search completed: April 22, 2004, 15:25:21
Job time : 1793 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 22, 2004, 17:06:22 ; Search time 2521 Seconds
(without alignments)
1599.125 Million cell updates/sec

Title: US-10-005-907-2
Perfect score: 715
Sequence: 1 MGNVLLKSLCLGENQKPK.....SVSRPCSTHEHDYEVVFP 135
Scoring table: BLOSOM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0
Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578
Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DEV=xlp
-O=/cgn2 1/USPTO.spool_p/US10005907/runat 21042004 113754 11441/app query.fasta_1.327
-DB=EST -OPT=fastap -SUFFIX=rat -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40 cdi -LIST=45
-DOALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907 @cgn 1.1 4237 @runat 21042004 113754 11441 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hcc:*
9: gb_est1:*
10: gb_est2:*
11: gb_hcc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_phg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	527	10	BF242113 601880401
2	173.5	24.3	1045	10	BF797507 602257174
3	169.5	23.7	924	12	BG389415 602414106
4	169.5	23.7	987	9	AL560376 602414106
5	169.5	23.7	1039	12	BM479887 AGENCOURT
6	169.5	23.7	1095	12	BG106563 602290373
7	169.5	23.7	1124	10	BF795688 602259560
8	166.5	23.3	1000	12	BM904106 AGENCOURT
9	166.5	23.3	1070	12	BM456595 AGENCOURT
10	166.5	23.3	1193	12	BM455198 AGENCOURT
11	159.5	22.3	1085	12	BM455407 AGENCOURT
12	156.5	21.9	508	29	AY404586 Homo sapi
13	156.5	21.9	683	10	AW963366 EST375439
14	156.5	21.9	876	13	BQ214408 AGENCOURT
15	150.5	21.0	1041	10	BG024891 602275814
16	149	20.8	1165	12	BM458339 AGENCOURT
17	146.5	20.5	1059	12	BM459235 AGENCOURT
18	145	20.3	495	9	AA311038 EST181808
19	142	19.9	526	14	CB286394 CMD36_D01
20	133.5	18.7	522	12	BG149093 uu88h07.Y
21	128	17.9	825	14	CB320849 AGENCOURT
22	127	17.8	520	12	BG145241 uu69e05.Y
23	127	17.8	524	12	BG276802 uv07d02.Y
24	127	17.8	526	12	BG276546 uv06d06.Y
25	125	17.5	661	10	BB629482 BB629482
26	125	17.5	3317	11	AK036553 Mus muscu
27	122	17.1	560	13	BU696543 L12in1317
28	119	16.6	884	13	BU390865 BX390865
29	118.5	16.6	815	13	BU92945 AGENCOURT
30	116.5	16.3	998	12	BG176880 602313443
31	112	15.7	479	13	BU429114 UI-HF-BNO
32	110.5	15.5	539	12	BG276071 uu89910.Y
33	109.5	15.3	538	10	AW519547 uu35b07.Y
34	108	15.1	1015	12	BM477520 AGENCOURT
35	105.5	14.8	1527	29	AG032114 Pan trogl
36	104.5	14.6	502	29	AY404588 Mus muscu
37	104	14.5	591	10	BF797042 602258082
38	103.5	14.5	927	12	BG177483 602314133
39	102.5	14.3	982	13	BX391573 BX391573
40	101.5	14.2	288	10	AW480906 33516 MAR
41	101.5	14.2	875	29	AG181241 Pan trogl
42	100.5	14.1	508	29	CE686854 tigr-gss
43	100	14.0	594	28	BH08248 RPT-24-3
44	99.5	13.9	1242	12	E1834393 603084439
45	99.5	13.9	1361	12	BG722475 602693720

ALIGNMENTS

RESULT 1
BF242113
LOCUS 601880401Fl NIH_MGC_55 Homo sapiens cDNA clone IMAGE:410906 5',
DEFINITION mRNA sequence.
ACCESSION BF242113.1 GI:11156040
VERSION BF242113
KEYWORDS EST.
SOURCE Homo sapiens
ORGANISM Homo sapiens
REFERENCE 1 (bases 1 to 527)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 cDNA Library Preparation: CLONETECH Laboratories, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LICM993 row: a column: 01
 High quality sequence stop: 518.

FEATURES
 source
 1. .527
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4109064"
 /tissue_type="from acute myelogenous leukemia"
 /lab_host="DH10B (TI phage-resistant)"
 /clone_lib="NIH_MGC_55"
 /note="Organ: bone marrow; Vector: pDNR-LIB (Clontech); Site 1: SfII (ggcgctcgcc); Site 2: SfiI (ggcattatggcc); Double-stranded cDNA was prepared from cell line RNA. 5' and 3' adaptors were used in cloning as follows: 5' adaptor sequence: 5'-CACGGCCATTATGGCC-3' and 3' adaptor sequence: 5'-ATTCTAGAGCGGAGGCGCCGATG-dT(30)BN-3', (where B = A, C, or G and N = A, C, G, or T). Average insert size 1.65 kb (range 0.9-4.0 kb). 14/15 colonies contained inserts by PCR. This library was enriched for full-length clones and was constructed by Clontech Laboratories (Palo Alto, CA)."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,66e-63 Length: 527
 Score: 715.00 Matches: 135
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 10 Gaps: 0

US-10-005-907-2 (1-135) x BF242113 (1-527)

QY 1 MetGlyAenTyLeuAryLysLeuSerCysLeuGlyGluAenGlnLysLysProLys 20
 DB 45 ATGGGAATTATCTCTCGAATACTGATGCTGCTGGAGAGATCAAGAAGACCCCAAG 104
 QY 21 LysGlyAenProAaspGluAryLysArgGlnGluMetThrPheGluAryLysLeu 40
 DB 105 AAAGGAACCCAGATGAGGAAGAAACGGCAGGAATGACTACATTGAAAGAAACTT 164
 QY 41 GlnAaspGlnApyLysSerGlnGluValSerThrSerAenGlnGluAenGluAen 60
 DB 165 CAAGATCAAGATAGAAAGACCAAGAGTTTCATCCACTTCTATCAGAAACAGAAAT 224
 QY 61 GlySerGlySerGluGluValCysTyThrValLeuAenHisfileProHisGlnArgSer 80
 DB 225 GGCAGTGGTCTCTGAAGAGTGTGTACACTGTCTATTATACATCCATCCCATCAGATCC 284
 QY 81 SerLeuSerSerAenAaspGlyTyrcGluAenIleAepSerLeuThrArgLysValArg 100
 DB 285 TCCCTGAGCTCAATGATGAGTGTATGAGACATTTGCTCCCTCAGCAAGAAAGTGAGA 344
 QY 101 GlnPheArgGluArySerGluThrGluTyThrValLeuAenHisfileProHisGlnArgPro 120
 DB 345 CAGTTTACAGAAAGGTACAGACAGAAATATGCCCTTCTTAGGACTTCTGTAGTAGGCT 404
 QY 121 CysSerCysThrHisGluHisAaspTyrcGluValValPheProHis 135

Db 405 TGTCTCTGACCCCATGAGCATGATTATGAGTGTGTGTTTCCACAC 449
 RESULT 2
 BF797507
 LOCUS
 DEFINITION 1045 bp mRNA linear EST 12-JAN-2001
 602257174F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4340323 5',
 mRNA sequence.
 BF797507
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
 REFERENCE
 1. (bases 1 to 1045)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM9952 row: d column: 20
 High quality sequence stop: 687.

FEATURES
 source
 1. .1045
 Location/Qualifiers
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:4340323"
 /tissue_type="lymphoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_85"
 /note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 1.867 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."

ORIGIN
 Alignment Scores:
 Pred. No.: 1.98e-07 Length: 1045
 Score: 173.50 Matches: 48
 Percent Similarity: 51.61% Conservative: 32
 Best Local Similarity: 30.97% Mismatches: 54
 Query Match: 24.27% Indels: 21
 DB: 10 Gaps: 5

US-10-005-907-2 (1-135) x BF797507 (1-1045)

QY 1 MetGlyAenTyLeuAryLysLeuSerCysLeuGlyGluAenGlnLysLysPro 19
 DB 103 ATGGGAATTCTCTGCTGAGAGAAACAGCGGCGAGACACTCAGAGATGCTTGG 162
 QY 20 -----LysLysGlyAenProAaspGluAryLysArg----- 30
 DB 163 AATGTGAGAATCAAAAGCCCCCAACAGAGAACATCCAGATGCTGGGATCATCATGCT 222
 QY 31 -----GlnGluMetThrPheGluAryLysLeuGlnAasp 42
 DB 223 GAAGGTGTTTCTGCTTCCATGGAAAAAATACTATCTTTTGAAGAAGG---CAAGAT 279
 QY 43 GlnApyLysLysSerGlnGluValSerThrSerAenGlnGluAenGlnGlySer 62
 DB 280 TCCCAAAACAAAATGAAGAATGTCTACTCTCCATCCAGGACAAATGTTCCACGACC 339
 QY 63 GlySerGluGluValCysTyThrValLeuAenHis---IleProHisGlnArgSer 81

Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 9 Gaps: 5

US-10-005-907-2 (1-135) x AL560376 (1-987)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
DB 179 ATGGGAAATTCCTGCTGAGAGAAAACAGCGGACAGAGAACTCAAGAGATGCTTGG 238
QY 20 -----LysLysGlyAsnProAspGluGlyArg----- 30
DB 239 AATGTAGAAATCAAGAACCCCAACAGAGAACATCCAGATGCTGGATCACCATTCCT 298
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 299 GAAGGCTTTCTGCTTCATGCAAAAAAATACATCTTTTGAAGAGG---CAAGAT 355
QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSer 62
DB 356 TCCCAAAACGAAATGAAAGATGATCATCTCCATCCAGAGAAATGTTACACAGACC 415
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
DB 416 TACTCAGAGAGAGCTGCTGCTTACCTCATCATCATCGGCTTCTGTACAGGCCATCA 475
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
DB 476 GGGAAATCTGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 535
QY 102 PheArgGluArgSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSer 121
DB 102 PheArgGluArgSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSer 121
QY 536 TCCCTGGAGAGAGAGCTGCTGCTTACCTCATCATCATCGGCTTCTGTACAGGCCATCA 595
DB 536 TCCCTGGAGAGAGAGCTGCTGCTTACCTCATCATCATCGGCTTCTGTACAGGCCATCA 595
QY 122 SerCys---ThrHisGlnHisAspTyrGluValValPheProHis 135
DB 596 CATGCCCGATCCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 640

RESULT 5
BM479887 1039 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT 6464830 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:557364
DEFINITION 5', mRNA sequence.
ACCESSION BM479887
VERSION BM479887.1 GI:18528929
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT 1 (bases 1 to 1039)
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strauberg, Ph.D.
Email: cga@pds-remail.nih.gov
Tissue Procurement: ATCC
CDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM12330 Row: 1 Column: 05
High quality sequence stop: 665.
Location/Qualifiers
1. 1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:557364"

FEATURES

1. 1039
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:557364"

/issue_type="embryonal carcinoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_92"
/note="Organ: testis; Vector: pCMV-Sport6; Site: 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 2.5 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC library."

ORIGIN

Alignment Scores:

Pred. No.: 5,02e-07 Length: 1039
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
DB: 12 Gaps: 5

US-10-005-907-2 (1-135) x BM479887 (1-1039)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
DB 81 ATGGGAAATTCCTGCTGAGAGAAAACAGCGGACAGAGAACTCAAGAGATGCTTGG 140
QY 20 -----LysLysGlyAsnProAspGluGlyArg----- 30
DB 141 AATGTAGAAATCAAGAACCCCAACAGAGAACATCCAGATGCTGGATCACCATTCCT 200
QY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 201 GAAGGCTTTCTGCTTCATGCAAAAAAATACATCTTTTGAAGAGG---CAAGAT 257
QY 43 GlnAspLysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSer 62
DB 258 TCCCAAAACGAAATGAAAGATGATCATCTCCATCCAGAGAAATGTTACACAGACC 317
QY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
DB 318 TACTCAGAGAGAGCTGCTGCTTACCTCATCATCATCGGCTTCTGTACAGGCCATCA 377
QY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
DB 378 GGGAACTTCTGCTGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 437
QY 102 PheArgGluArgSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSer 121
DB 438 TCCCTGGAGAGAGAGCTGCTGCTTACCTCATCATCATCGGCTTCTGTACAGGCCATCA 497
QY 122 SerCys---ThrHisGlnHisAspTyrGluValValPheProHis 135
DB 498 CATGCCCGATCCCGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATG 542

RESULT 6
BG106563 1095 bp mRNA linear EST 30-JAN-2001
LOCUS 602290373P1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4385178 5',
DEFINITION mRNA sequence.
ACCESSION BG106563
VERSION BG106563.1 GI:12600409
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

REFERENCE
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
TITLE Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
JOURNAL NIH-MGC http://mgi.nci.nih.gov/
COMMENT 1 (bases 1 to 1095)
Unpublished (1999)
National Institutes of Health, Mammalian Gene Collection (MGC)
Contact: Robert Strauberg, Ph.D.
Email: cga@pds-remail.nih.gov
Tissue Procurement: Louis Straub, M.D., Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM10066 row: a column: 19
 High quality sequence stop: 641.

FEATURES

source

Location/Qualifiers

```
1..1095
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4385178"
/cisue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_85"
/Note="Organ: lymph; Vector: PCMV-SPORE; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:

```
Pred. No.: 5.32e-07 Length: 1095
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
Gaps: 5
```

US-10-005-907-2 (1-135) x BGI06563 (1-1095)

```
OY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLeuPro--- 19
DB 43 ATGGGAATTTCTGCTGAGAGAAACAGCGGCGAGCAAGCACTCAAGAGATGCTTGG 102
OY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
DB 103 AATGTGAATGCAAGACCCCAACAGAGAACATCCAGATCGGGATCACCATTGCGCT 162
OY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 163 GAAGGTTGTTCTGCTTCATGGAAGAAATCACTATTGTAAGAGAG---CAAGAT 219
OY 43 GlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
DB 220 TCCCAAAACGAAAGAAATGATCACTACTCCATCCAGAGCAATGTTGACAGAC 279
OY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
DB 280 TACTAGAGAGAGCTGTGCTATACCTCATCAATCACTCGGTTCTCTGTACAGGCCATCA 339
OY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
DB 340 GGAAGACTGCTGCAAGAGACTGATGAAATGTTCCCTCCAGAAAGCAGAGAGCCAGAGAG 399
OY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
DB 400 TCCTTGAGAGAGACTGAGACTGAGATTCATCTTCAATATAGCCCTTCTACAGAGCCAGG 459
OY 122 SerCys---ThriSgIuHisAspTyrGluValValPheProHis 135
DB 460 CATGCCCATCCCCAGAGAGATGAAATATGAACTTCTCATGCTCAGC 504
```

RESULT 7
 BF795688 1124 bp mRNA linear EST 12-JAN-2001
 LOCUS 602259560F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:4342719 5'
 DEFINITION mRNA sequence.
 ACCESSION BF795688
 VERSION BF795688.1 GI:12100659
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

1 (bases 1 to 1124)

AUTHORS

NIH-MGC <http://mgi.nci.nih.gov/>.

TITLE

National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL

Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
 Email: cgabbs-remail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L1AM9958 row: h column: 16
 High quality sequence stop: 715.

FEATURES

source

Location/Qualifiers

```
1..1124
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4342719"
/cisue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_id="NIH_MGC_85"
/Note="Organ: lymph; Vector: PCMV-SPORE; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."
```

ORIGIN

Alignment Scores:

```
Pred. No.: 5.48e-07 Length: 1124
Score: 169.50 Matches: 47
Percent Similarity: 51.61% Conservative: 33
Best Local Similarity: 30.32% Mismatches: 54
Query Match: 23.71% Indels: 21
Gaps: 5
```

US-10-005-907-2 (1-135) x BF795688 (1-1124)

```
OY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLeuPro--- 19
DB 118 ATGGGAATTTCTGCTGAGAGAAACAGCGGCGAGCAAGCACTCAAGAGATGCTTGG 177
OY 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
DB 178 AATGTGAATGCAAGACCCCAACAGAGAACATCCAGATCGGGATCACCATTGCGCT 237
OY 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
DB 238 GAAGGTTGTTCTGCTTCATGGAAGAAATCACTATTGTAAGAGAG---CAAGAT 294
OY 43 GlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
DB 295 TCCCAAAACGAAAGAAATGATGAAATGTTCCCTCCAGAAAGCAGAGAGCCAGAGAG 354
OY 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
DB 355 TACTAGAGAGAGCTGTGCTATACCTCATCAATCACTCGGTTCTCTGTACAGGCCATCA 414
OY 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
DB 415 GGAAGACTGCTGCAAGAGACTGATGAAATGTTCCCTCCAGAAAGCAGAGAGCCAGAGAG 474
OY 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
DB 475 TCCTTGAGAGAGACTGAGACTGAGATTCATCTTCAATATAGCCCTTCTACAGAGCCAGG 534
OY 122 SerCys---ThriSgIuHisAspTyrGluValValPheProHis 135
```



```

Db      196  CCTTGGAAATGAGAAATGCAAGAGCCCAACAGAGAACATCCAGATCGCTGGATCACCAT 255
QY      31  -----GlnGluMetThrThrPheGluArgLysLeu 40
Db      256  ATGCTGAAAGGGTGTCTCCCTCCATGATGAAATAATCTCATTTTGAAGAGG--- 312
QY      41  GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsn 60
Db      313  CAAGATTCCCAAAACGAAATGAAGAAATGCTACTTCTCCATCCAGGACAAATGTTGAC 372
QY      61  GlySerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArg 79
Db      373  CAGACCTTACTCAGAGGAGCTGTGCTATACCTCATCAATCATCGGTTCTCTTACAGG 432
QY      80  SerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99
Db      433  CCATCAGGGAACCTCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGCTGAGAGACC 492
QY      100  ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db      493  AGAGAGTCCTTGGGAGGAACCTGAGACTGAGTATTCACCTTACATATGCTTCTACAGAC 552
QY      120  ProCysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db      553  CCCAGGATGCCCATCCCAAGAGATGAATGAACTTCTCATGCTCTCAC 603

RESULT 10
BM455198 1193 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6405645 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500141
DEFINITION 5', mRNA sequence.
ACCESSION BM455198
VERSION BM455198.1 GI:18504238
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1193)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12134 row: j column: 14
High quality sequence stop: 592.
Location/Qualifiers
1..1193
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500141"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

FEATURES
source
1..1193
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500141"
/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH_MGC_85"
/note="Organ: lymph; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally; oligo-dT primed.
Average insert size 1.867 Kb. Library enriched for
full-length clones and constructed by Life Technologies.
Note: this is a NIH_MGC Library."

ORIGIN
Alignment Scores: 1.18e-06 Length: 1193
Pred. No.: 166.50 Matches: 48
Score:

```

```

Percent Similarity: 52.23% Conservative: 34
Best Local Similarity: 30.57% Mismatches: 52
Query Match: 23.29% Indels: 23
DB: 12 Gaps: 6

US-10-005-907-2 (1-135) x BM455198 (1-1193)

QY      1  MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGlu-----AsnGlnLysLys 18
Db      65  ATGGGAAATTTCTGCTGAGAGAAACAGAGTTTCAGGCGGCGAGACACACATCAAGAGATG 124
QY      19  Pro-----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db      125  CTTTGGAAATGTGAGAAATCCAAAGCCCCCAACAGAGAACATCCAGATCGTGGATCACCAT 184
QY      31  -----GlnGluMetThrThrPheGluArgLysLeu 40
Db      185  ATGCTGTAAGGGTGTCTCTGCTTCCATGGAATAATACTCATTTTGAAGAGG--- 241
QY      41  GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
Db      242  CAAGATTCCCAAAACGAAATGAAGAAATGCTACTTCTCCATCCAGGACAAATGTTGAC 301
QY      61  GlySerGlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArg 79
Db      302  CAGACCTTACTCAGAGGAGCTGTGCTATACCTCATCAATCATCGGTTCTCTGTACAAG 361
QY      80  SerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal 99
Db      362  CCATCAGGGAACCTCTGCTGAAGAGTACTATGAGAAATGTTCCCTGCAAGCTGAGAGACC 421
QY      100  ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db      422  AGAGAGTCCTTGGGAGGAACCTGAGACTGAGTATTCACCTTACATATGCTTCTACAGAC 481
QY      120  ProCysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db      482  CCCAGGATGCCCATCCCAAGAGATGAATGAACTTCTCATGCTCTCAC 532

RESULT 11
BM455407 1085 bp mRNA linear EST 05-FEB-2002
LOCUS AGENCOURT_6407334 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:5500391
DEFINITION 5', mRNA sequence.
ACCESSION BM455407
VERSION BM455407.1 GI:18504436
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1085)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: Lou Staudt
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM12135 row: d column: 24
High quality sequence start: 19
High quality sequence stop: 760.
Location/Qualifiers
1..1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500391"

FEATURES
source
1..1085
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5500391"

```

```

/tissue_type="lymphoma, cell line"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NIH MGc 85"
/note="Organ: lymph; Vector: pGMW-
Site_2: SalI; Cloned unidirectional.
Average insert size 1.867 Kb. Library
full-length clones and constructed
Note: this is a NIH MGc Library."

```

COMMENT This sequence was made by sequencing genomic exons and ordering them based on alignment.

FEATURES

```
/organism="Homo sapiens"  
/mol_type="genomic DNA"  
/db_xref="taxon:9606"
```

gene

ORIGIN

Pred. No.:	5.49e-06	Length:	1085
Score:	159.50	Matches:	37
Percent Similarity:	60.75%	Conservative:	28
Best Local Similarity:	34.58%	Mismatches:	39
Query Match:	22.31%	Indels:	23
DB:	12	Gaps:	3

US-10-005-907-2 (1-135) x BM455407 (1-1085)

31	QY	GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal	50
439	Db	AAAAAATACTCATTTTTGAAGAGG---CAAGATTCCCAAAACGAAATGAAGAATG	495
51	QY	SerSerThrSerAsnGlnGluAsnGlySerGlySerGluGluValCysTyrThr	70
496	Db	TCATCTACTCCCATCCAGACAACTGTGACCAGACTTACAGAGAGCTGTCTATAACC	555
71	QY	ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyr	89
556	Db	CTCATCAATCATCGGGTTCCTGTGTACAGGCCATCANGGAACCTGCTGAAGAGTACTAT	615
90	QY	GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu	109
616	Db	GAGAATGTCCTCGAAAGCTGAGAGACCCAGAGAGCTCTTGGAGGAACTGAGACTGAG	675
110	QY	TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp	128
676	Db	TATTCACCTTCTACATATGCTTCTTACAGACCCAGGCCATGCCGATCCCAAGAGATGAA	735
129	QY	TyrGluValValPheProHis	135
736	Db	TATGAACCTTCTCATGCTTCAC	756

RESULT 12	AY404586	508 bp	DNA	linear	GSS 16-DEC-2003
LOCUS	AY404586				
DEFINITION	Homo sapiens GGET2 gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.				

ACCESSION
AY404586
VERSION
AY404586.1
KEYWORDS
GSS.

SOURCE	Homo sapiens (human)
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

ORGANISM	Homo sapiens
1	
2	
3	
4	
5	
6	
7	
8	
9	
10	
11	
12	
13	
14	
15	
16	
17	
18	
19	
20	
21	
22	
23	
24	
25	
26	
27	
28	
29	
30	
31	
32	
33	
34	
35	
36	
37	
38	
39	
40	
41	
42	
43	
44	
45	
46	
47	
48	
49	
50	
51	
52	
53	
54	
55	
56	
57	
58	
59	
60	
61	
62	
63	
64	
65	
66	
67	
68	
69	
70	
71	
72	
73	
74	
75	
76	
77	
78	
79	
80	
81	
82	
83	
84	
85	
86	
87	
88	
89	
90	
91	
92	
93	
94	
95	
96	
97	
98	
99	
100	

REFERENCE
AUTHORS

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 508)

Clark, A.G., Gnanowski, S., Nielson, R., Thomas, P., Kejarinwal, A.,
Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
Adams, M.D. and Cargill, M.

TITLE	JOURNAL
Inferring nonneutral evolution from human-chimp-mouse orthologous gene trios	Science 302 (5652): 1960-1963 (2003)

PUBMED 14671302
2 (bases 1 to 508)
REFERENCE
AUTHORS
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarawal, A.,
Todd, M.A., Tanenbaum, D.M., Civallo, D.R., Lu, F., Murphy, B.,
Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Sainsky, J.J.,
Adams, M.D. and Carcilli, M.

TITLE
Direct Submission
JOURNAL
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA

Alignment Scores:	4.72e-06	Length:	508
Pred. No.:	186.50	Matches:	37
Score:	186.50	Conservative:	28
Percent Similarity:	60.75%	Mismatches:	39
Best Local Similarity:	34.5%	Indels:	3
Query Match:	21.89%	Gaps:	3
DB:	29		

US-10-005-907-2 (1-135) x AY404586 (1-508)

Qy	31	GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal	50
Db	116	AAAAAATATCATCTTTTGAAGAGG--CAAGATCCCAAAACGAANAATGAAGAATG	172
Qy	51	SerSerThrSerAsnGlnGluAsnGluSerGlySerGluGluValCysThrThr	70
Db	173	TCATCATCTCCATCCAGGACAATGTTCACACACTTCTCAGAGAGCTGTGCTATACC	232
Qy	71	ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyr	89
Db	233	CTCATCAATCATCGGGTTCTCTGTACAAAGCCATCAAGGAATCTCTCTGAAGAGTACTAT	292
Qy	90	GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu	109
Db	293	GAGAATGTTCCCTGCAAAAGCTGAGAGACCAGAGAGTCCCTTGGAGGAAGACTCAGACTGAG	352
Qy	110	TyrAlaLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp	128
Db	353	TATTCATCTTCTACATATGCTTCTTACAGACCCAGAGCATGCCGATGCCAGAGAAGATGAA	412

RESULT 13

AW963366

LOCUS

DEFINITION EST375439 MAGE resequences, MAGE Homo sapiens cDNA, mRNA sequence.

ACCESSION AW96338

VERSION AW963366.1 GI:8153202

KEYWORDS

SOURCE Homo sapiens (human)

ORGANISM

Eukaryota; Metazoa; Chordata; Cranialata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE

AUTHORS

Holt, I.E., Saeed, A.I., Sharov, V., Lee, N.H., Yeatman, T.J. and Quackenbush, J.

TITLE Assessment of gene expression patterns in a model of colon tumor metastasis using a 19,200 element cDNA microarray
JOURNAL Unpublished (2000)

COMMENTS

The Institute for

9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 3528
Fax: 301 838 0208
Email: john@tigr.org
Plate: 187

Seq primer: Reverse.

FEATURES	Location/Qualifiers
1. <i>the</i>	1. <i>the</i> (100%)
2. <i>the</i>	2. <i>the</i> (100%)
3. <i>the</i>	3. <i>the</i> (100%)
4. <i>the</i>	4. <i>the</i> (100%)
5. <i>the</i>	5. <i>the</i> (100%)
6. <i>the</i>	6. <i>the</i> (100%)
7. <i>the</i>	7. <i>the</i> (100%)
8. <i>the</i>	8. <i>the</i> (100%)
9. <i>the</i>	9. <i>the</i> (100%)
10. <i>the</i>	10. <i>the</i> (100%)
11. <i>the</i>	11. <i>the</i> (100%)
12. <i>the</i>	12. <i>the</i> (100%)
13. <i>the</i>	13. <i>the</i> (100%)
14. <i>the</i>	14. <i>the</i> (100%)
15. <i>the</i>	15. <i>the</i> (100%)
16. <i>the</i>	16. <i>the</i> (100%)
17. <i>the</i>	17. <i>the</i> (100%)
18. <i>the</i>	18. <i>the</i> (100%)
19. <i>the</i>	19. <i>the</i> (100%)
20. <i>the</i>	20. <i>the</i> (100%)
21. <i>the</i>	21. <i>the</i> (100%)
22. <i>the</i>	22. <i>the</i> (100%)
23. <i>the</i>	23. <i>the</i> (100%)
24. <i>the</i>	24. <i>the</i> (100%)
25. <i>the</i>	25. <i>the</i> (100%)
26. <i>the</i>	26. <i>the</i> (100%)
27. <i>the</i>	27. <i>the</i> (100%)
28. <i>the</i>	28. <i>the</i> (100%)
29. <i>the</i>	29. <i>the</i> (100%)
30. <i>the</i>	30. <i>the</i> (100%)
31. <i>the</i>	31. <i>the</i> (100%)
32. <i>the</i>	32. <i>the</i> (100%)
33. <i>the</i>	33. <i>the</i> (100%)
34. <i>the</i>	34. <i>the</i> (100%)
35. <i>the</i>	35. <i>the</i> (100%)
36. <i>the</i>	36. <i>the</i> (100%)
37. <i>the</i>	37. <i>the</i> (100%)
38. <i>the</i>	38. <i>the</i> (100%)
39. <i>the</i>	39. <i>the</i> (100%)
40. <i>the</i>	40. <i>the</i> (100%)
41. <i>the</i>	41. <i>the</i> (100%)
42. <i>the</i>	42. <i>the</i> (100%)
43. <i>the</i>	43. <i>the</i> (100%)
44. <i>the</i>	44. <i>the</i> (100%)
45. <i>the</i>	45. <i>the</i> (100%)
46. <i>the</i>	46. <i>the</i> (100%)
47. <i>the</i>	47. <i>the</i> (100%)
48. <i>the</i>	48. <i>the</i> (100%)
49. <i>the</i>	49. <i>the</i> (100%)
50. <i>the</i>	50. <i>the</i> (100%)
51. <i>the</i>	51. <i>the</i> (100%)
52. <i>the</i>	52. <i>the</i> (100%)
53. <i>the</i>	53. <i>the</i> (100%)
54. <i>the</i>	54. <i>the</i> (100%)
55. <i>the</i>	55. <i>the</i> (100%)
56. <i>the</i>	56. <i>the</i> (100%)
57. <i>the</i>	57. <i>the</i> (100%)
58. <i>the</i>	58. <i>the</i> (100%)
59. <i>the</i>	59. <i>the</i> (100%)
60. <i>the</i>	60. <i>the</i> (100%)
61. <i>the</i>	61. <i>the</i> (100%)
62. <i>the</i>	62. <i>the</i> (100%)
63. <i>the</i>	63. <i>the</i> (100%)
64. <i>the</i>	64. <i>the</i> (100%)
65. <i>the</i>	65. <i>the</i> (100%)
66. <i>the</i>	66. <i>the</i> (100%)
67. <i>the</i>	67. <i>the</i> (100%)
68. <i>the</i>	68. <i>the</i> (100%)
69. <i>the</i>	69. <i>the</i> (100%)
70. <i>the</i>	70. <i>the</i> (100%)
71. <i>the</i>	71. <i>the</i> (100%)
72. <i>the</i>	72. <i>the</i> (100%)
73. <i>the</i>	73. <i>the</i> (100%)
74. <i>the</i>	74. <i>the</i> (100%)
75. <i>the</i>	75. <i>the</i> (100%)
76. <i>the</i>	76. <i>the</i> (100%)
77. <i>the</i>	77. <i>the</i> (100%)
78. <i>the</i>	78. <i>the</i> (100%)
79. <i>the</i>	79. <i>the</i> (100%)
80. <i>the</i>	80. <i>the</i> (100%)
81. <i>the</i>	81. <i>the</i> (100%)
82. <i>the</i>	82. <i>the</i> (100%)
83. <i>the</i>	83. <i>the</i> (100%)
84. <i>the</i>	84. <i>the</i> (100%)
85. <i>the</i>	85. <i>the</i> (100%)
86. <i>the</i>	86. <i>the</i> (100%)
87. <i>the</i>	87. <i>the</i> (100%)
88. <i>the</i>	88. <i>the</i> (100%)
89. <i>the</i>	89. <i>the</i> (100%)
90. <i>the</i>	90. <i>the</i> (100%)
91. <i>the</i>	91. <i>the</i> (100%)
92. <i>the</i>	92. <i>the</i> (100%)
93. <i>the</i>	93. <i>the</i> (100%)
94. <i>the</i>	94. <i>the</i> (100%)
95. <i>the</i>	95. <i>the</i> (

FEATURES

source

1. .683
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone_lib="MAGE resequences, MAGE"
 /note="Vector: pBluescriptSKm"

ORIGIN

Alignment Scores:

Pred. No.: 6.59e-06 Length: 683
 Score: 156.50 Matches: 37
 Percent Similarity: 60.75% Conservative: 28
 Best Local Similarity: 34.58% Mismatches: 39
 Query Match: 21.89% Indels: 3
 DB: 10 Gaps: 3

US-10-005-907-2 (1-135) x AW963366 (1-683)

QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal 50
 DB 57 AAAAAAATACCTCATTTTGAAGAGAGG--CAAGATTCCCAAAACGAAATGAAGAATG 113
 QY 51 SerSerThrSerAsnGlnGluAsnGlySerGlySerGlnGluValCysTyrThr 70
 DB 114 TCATCTACTCCATCCAGGACATGTTGACACACCTACTCAGAGGAGCTGTCTATACC 173
 QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyr 89
 DB 174 CTCATCAATCATCGGTTCTCTGTACAAAGCCATCAGGAACTCTGCTGAAGAGTACTAT 233
 QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
 DB 234 GAGAATGTTCCCTCAAGCTGAAGACCCAGAGAGTCTCTGGAGGAAGTGAAGATGAG 293
 QY 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128
 DB 294 TATTCATCTTCTACATATGCTCTTACAGACCCAGGATGCCGATCCCGAAGATGAA 353
 QY 129 TyrGluValValPheProHis 135
 DB 354 TATGAATCTTCTCATGCTCTCAC 374

RESULT 14
 BQ214406
 LOCUS
 DEFINITION AGENCOURT 7592611 NIH_MGC_92 Homo sapiens cDNA clone IMAGE:6070069
 5', mRNA sequence.

ACCESSION BQ214406
 VERSION BQ214406.1 GI:20395806
 KEYWORDS EST.
 SOURCE Homo sapiens (human)

ORGANISM
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE
 1 (bases 1 to 876)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

AUTHORS
 TITLE
 JOURNAL
 COMMENT
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: ATCC
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Agencourt Bioscience Corporation
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LLAM1353 row: m column: 14
 High quality sequence stop: 678.

FEATURES

source

1. .876
 /organism="Homo sapiens"
 /mol_type="mRNA"

/db_xref="taxon:9606"
 /clone="IMAGE:6070069"
 /tissue_type="embryonal carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_92"
 /note="Organ: testis; Vector: pCMV-SPORT6; Site 1: NOL1;
 Site 2: Sall; Cloned unidirectionally; oligo-dT primed.
 Average insert size 2.5 kb. Library enriched for
 full-length clones and constructed by Life Technologies.
 Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 8.72e-06 Length: 876
 Score: 156.50 Matches: 37
 Percent Similarity: 60.75% Conservative: 28
 Best Local Similarity: 34.58% Mismatches: 39
 Query Match: 21.89% Indels: 3
 DB: 13 Gaps: 3

US-10-005-907-2 (1-135) x BQ214406 (1-876)

QY 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal 50
 DB 384 AAAAAAATACCTCATTTTGAAGAGG--CAAGATTCCCAAAACGAAATGAAGAATG 440
 QY 51 SerSerThrSerAsnGlnGluAsnGlySerGlySerGlnGluValCysTyrThr 70
 DB 441 TCATCTACTCCATCCAGGACATGTTGACACACCTACTCAGAGGAGCTGTCTATACC 500
 QY 71 ValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyr 89
 DB 501 CTCATCATCTCATCGGTTCTCTGTACAAAGCCATCAGGAACTCTCTGTAAGAGTACTAT 560
 QY 90 GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGlu 109
 DB 561 GAGAATGTTCCCTCAAGCTGAAGACCCAGAGAGTCTCTGGAGGAAGTGAAGATGAG 620
 QY 110 TyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHisAsp 128
 DB 621 TATTCATCTTCTACATATGCTCTTACAGACCCAGGATGCCGATCCCGAAGATGAA 680
 QY 129 TyrGluValValPheProHis 135
 DB 681 TATGAATCTTCTCATGCTCTCAC 701

RESULT 15

BQ24891

LOCUS

DEFINITION

BQ24891 1041 bp mRNA linear EST 24-JAN-2001
 602275814F1 NIH_MGC_85 Homo sapiens cDNA clone IMAGE:436899 5',
 mRNA sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1041)
 NIH-MGC http://mgi.nci.nih.gov/.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)

Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Louis Staudt, M.D., Ph.D.
 CDNA Library Preparation: Life Technologies, Inc.
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov

Plate: LLAM10010 row: k column: 04
 High quality sequence start: 2

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 22, 2004, 16:07:56 ; Search time 349 Seconds

(without alignments)
1643.286 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715

Sequence: 1 MGNYLRLKLSCLGKQKPK.....SVSRPCSTHEHDYEVVFPFH 135

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2.1/USPTO.spool_p/US10005907/runat.21042004.113753.11422/app_query.fasta.1.327
-DB=N Geneseq 29Jan04 -QMT=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US10005907 -SCGN.1.1.708 -runat.21042004.113753.11422 -NCPU=6 -ICPU=3
-NO MMAP -LARGESQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEVTIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPEXT=10 -XGAPEXT=0.5 -FGAPOPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELEXT=7

Database :

N Geneseq 29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001as:*

5: Geneseq2001bs:*

6: Geneseq2002as:*

7: Geneseq2003as:*

8: Geneseq2003bs:*

9: Geneseq2003cs:*

10: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	715	100.0	667	9	ADC30760 Human nov
2	715	100.0	3762	6	ABN81319 Human mas
3	169.5	23.7	732	7	ACC59951 Human IRA
4	159	22.2	1643	6	ABL99791 Human sec
5	149	20.8	775	7	ABX05108 Human nov
6	98	13.7	1892	2	AAZ24890 Human sec
7	98	13.7	1892	7	ADA56130 Gene enco
8	98	13.7	1892	7	ADA39940 Human sec

9	98	13.7	1892	9	ADC73578
10	98	13.7	1892	9	ADD37623
11	96.5	13.5	584	9	ADE62022
12	96	13.4	2301	3	AAV20445
13	96	13.4	2301	3	AAA09303
14	93	13.0	1824	4	ABL27495
15	92	12.9	654	5	AA571153
16	92	12.9	654	5	AA575467
17	92	12.9	654	5	AA569553
18	91.5	12.8	3182	5	ABX71302
19	91.5	12.8	7488	6	ABL33112
20	90	12.6	1671	4	AAH99858
21	90	12.6	1899	2	AAH06785
22	90	12.6	1912	3	AAZ50927
23	90	12.6	7992	6	ABK32895
24	90	12.6	8083	5	AA570692
25	90	12.6	8257	8	ACA90154
26	90	12.6	8493	7	ACA03930
27	90	12.6	8503	4	AA158253
28	90	12.6	8503	8	ADB48220
29	90	12.6	8527	4	AA160039
30	89.5	12.5	1027	2	AA584690
31	89.5	12.5	1027	6	ABK47580
32	89.5	12.5	1071	4	ABA83301
33	89.5	12.5	1071	8	ACH04802
34	89.5	12.5	1071	8	ACD44612
35	89.5	12.5	5695	3	AA576938
36	89.5	12.5	9415	6	ABL32295
37	89	12.4	864	7	ABX09875
38	89	12.4	902	6	ABQ46432
39	89	12.4	902	6	ABQ46433
40	89	12.4	13574	3	AAA81529
41	89	12.4	110000	3	AAF21613
42	89	12.4	172325	3	AAF21613
43	88.5	12.4	3430	8	ACF05253
44	88	12.3	7409	4	ABL13405
45	88	12.3	10215	4	ABL13404

ALIGNMENTS

RESULT 1

ADC30760

ID ADC30760 standard; cDNA; 667 BP.

XX ADC30760;

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

XX AC

Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
Halley-Vicente D, Drmanac RT;

WPI; 2003-371981/35.

P-PSDB; ADC31731.

New polynucleotide and polypeptide useful for diagnosing, preventing or treating conditions such as neurodegenerative diseases, anemias, platelet disorders, wounds, ulcers, osteoporosis, autoimmune diseases or cancer.

Claim 1; SEQ ID NO 842; 1185pp; English.

The invention relates to 971 novel human cDNA sequences (ADC29919-ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The invention also relates to nucleic acid sequences over 99% identical with the novel human cDNAs. The invention additionally encompasses expression vectors and host cells comprising a nucleic acid of the invention; the recombinant production of a polypeptide of the invention; an antibody against a polypeptide of the invention; a method of detecting polynucleotides or polypeptides of the invention; and methods of identifying a compound which binds to a polypeptide of the invention. The invention further discloses methods of preventing, treating or ameliorating a medical condition; kits comprising polynucleotide probes and/or monoclonal antibodies for carrying out the methods of the invention; methods for the identification of compounds that modulate the expression or activity of the polynucleotide and/or polypeptide; and 767 contig sequences corresponding to the cDNA sequences of the invention (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628-ADC33394). The nucleic acids and polypeptides of the invention are useful in diagnostics, drug screening, forensics, gene mapping, in the identification of mutations responsible for genetic disorders or other traits, for assessing biodiversity, and in producing many other types of data and products dependent on DNA and amino acid sequences. They are also used for treating diseases such as Parkinson's disease, Alzheimer's disease and other neurodegenerative diseases, anaemia, platelet disorders, wounds, ulcers, osteoporosis, autoimmune diseases or cancer. The nucleic acids may also be used as hybridisation probes or primers, and in the recombinant production of a protein. The polypeptides are also useful in generating antibodies, as molecular weight markers, and as food supplements. The present sequence represents a specifically claimed human cDNA sequence of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Alignment Scores:

Score: 8,248-76 Length: 667
Pred. No.: 715.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 9 Gaps: 0

US-10-005-907-2 (1-135) x ADC30760 (1-667)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLeuProlys 20
DB 186 ATGGGAAATATCTCTCGGAAACTCATGTTGCTGGGAGAGAAATCAAAAGAGCCCAAG 245
QY 21 LysGlyAsnProAspGlnGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 246 AAGGNAACCCAGATGAGGAGAAAGAAACCGGAGGAATGACTACATTGGAAGAAACTT 305
QY 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
DB 306 CAAGATCAAGATAAGAAACCAAGAAAGTTTCATCCACTTCTTAATCAGGAAACAGGAAT 365
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 366 GGCAGTGGTCTGAGAGAGTGGTCACTGCTGCTATATCATCATCCCATCCATCAGAGATCC 425

QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerSerLeuThrArgLysValArg 100
DB 426 TCCCTGAGCTCCAAATGATGATGGCTATGAGAACATTCGACTCCCTCACAGGAAGTGA 485
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 486 CAGTTTGAAGAAAGGTTCAGACAGAGATATGCCCTCTCTAGACTTCTGTTAGGCGCT 545
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 546 TGTCTCTGCACCCATGAGCATGATTATGAGTGTGTGTTCCACAC 590

RESULT 2

ABN81319
ID ABN81319 standard; cDNA; 3762 BP.
XX
AC ABN81319;
XX
DT 30-AUG-2002 (first entry)
XX
DE Human mast cell related gene MCL SEQ ID NO 1.
XX
KW Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;
KW gene; ss.
XX
OS Homo sapiens.
XX
XX Key Location/Qualifiers
FT CDS 25..432
FT /tag= a
FT /product= "MCL"
XX
PN WO200246389-A2.
XX
PD 13-JUN-2002.
XX
XX 07-DEC-2001; 2001WO-US046180.
XX
XX 08-DEC-2000; 2000US-0251835P.
PR 14-MAR-2001; 2001US-0275479P.
PR 28-MAR-2001; 2001US-0279115P.
PR 02-APR-2001; 2001US-0280143P.
XX
XX (UNIO) UCB SA.
XX
XX Nocka K, Pirozzi G, Einstein R;
XX
XX WPI; 2002-508560/54.
XX P-PSDB; ABB77569.
XX
XX Novel isolated nucleic acids that are differentially expressed in mast
XX cells in patients with allergic hypersensitivity, encoding proteins
XX associated with mast cell regranulation and allergic hypersensitivity.
XX
XX Claim 1; Page 95-97; 119pp; English.

The invention relates to isolated nucleic acid (ABN81319-ABN81324), corresponding to genes differentially expressed in mast cells following activation or in patients with allergic hypersensitivity disease, (I) that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of (II) if at least 6 amino acids. (II) is useful for identifying binding partners. (I) or (II) is useful for diagnosing or treating a disease state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma, urticaria or atopic dermatitis or mastocytosis) in a subject which involves determining the level of expression of (I) or (II). A computer system, comprising a database containing information identifying the expression level in a tissue or at least one mast cell of (I), is useful for presenting information to identify the relative expression level of (I). (II) is used as a marker to detect, diagnose or identify an allergic response in a patient. The protein can also serve as a target that modulate gene expression or activity and as an antigen to raise

CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity, is
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
 XX
 SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 8,19e-75 Length: 3762
 Score: 715.00 Matches: 135
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 6 Gaps: 0

US-10-005-907-2 (1-135) x ABN81319 (1-3762)

Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAenGlnLysProLys 20
 Db 25 ATGGGAAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAG 84
 Qy 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
 Db 85 AAAGGAACCCAGATGAGGAAGAAACGCGAGGAATGACTACATTTTGAAGAAACTT 144
 Qy 41 GlnAspGlnAspLysSerGlnGluValSerSerThrSerAenGlnGluAenGluAen 60
 Db 145 CAAGATCAAGATAGAAAGCCAGAAAGTTTCATCCACTTCAATCAGGAAACGAGAAAT 204
 Qy 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 Db 205 GGCAGTGGTCTTGAAGAGTGTGCTACATGTCATTAATCAATCCCTCCATCAGATCC 264
 Qy 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
 Db 265 TCCCTGAGTCCCAATCATGATGCTGATGATGATGATGATGATGATGATGATGATGAT 324
 Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
 Db 325 CAGTTTAGAAGAGGTGAGAGACAGATATGCCCCCTTCTTAGGACTTCTGTAGAGCCT 384
 Qy 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
 Db 385 TGTTCCTGCACCCATGAGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 429

RESULT 3

ACCS9951
 ID ACCS9951 standard; cDNA; 732 BP.

ACCS9951;

07-JUL-2003 (first entry)

Human IRAP-28 encoding cDNA SEQ ID NO:63.

Human; immune response associated protein; IRAP; anti-HIV; anti-allergic;
 anti-inflammatory; antianemic; antiparkinsonian; nootropic; antithyroid;
 anticonvulsant; antiarteriosclerotic; antidiabetic; immunosuppressive;
 cytototoxic; hepatotropic; dermatological; antidiabetic; nephrotropic;
 angiot; thyromimetic; neuroprotective; osteopathic; antiarthritic;
 antiparasitic; antihelminthic; antipsoriatic; uropathic; protozoacide;
 antirheumatic; haemostatic; antibacterial; virucide; ophthalmological;
 fungicide; gene therapy; immune system disorder; neurological disorder;
 developmental disorder; muscle disorder; cell proliferative disorder;
 infection; gene; ss.

Homo sapiens.

Key Location/Qualifiers
 134..670
 CDS /*tag= a

/product= "IRAP-28"

WO2003025542-A2.

27-MAR-2003.

19-SEP-2002; 2002WO-US029979.

21-SEP-2001; 2001US-0324034P.

05-OCT-2001; 2001US-0327395P.

12-OCT-2001; 2001US-0328923P.

09-OCT-2001; 2001US-0342810P.

09-NOV-2001; 2001US-0344468P.

21-NOV-2001; 2001US-0332140P.

07-DEC-2001; 2001US-0340282P.

09-JAN-2002; 2002US-0347693P.

20-FEB-2002; 2002US-0358279P.

01-MAR-2002; 2002US-0361088P.

15-MAR-2002; 2002US-0364494P.

10-MAY-2002; 2002US-0379876P.

11-JUN-2002; 2002US-0388180P.

(INCY-) INCYTE GENOMICS INC.

Ho A, Baughn MR, Becha SD, Burford N, Elliott VS, Emerling BM;

Forsythe U, Gorvad AE, Griffin JA, Hafalia AJA, Honchell CD;

Ison CH, Burrill JD, Blake JJ, Lai PG, Lee EA, Marquis JP;

Lehr-Mason PM, Lee S, Sprague WM, Swarnakar A, Tang YT, Tran B;

Tran UK, Bhatia U, Chawla NK, Warren BA, Zheng W, Xu Y, Yue H;

WPI; 2003-363161/34.

P-PSDB; ABR43232.

New human immune response associated proteins and polynucleotides, useful
 for diagnosing, treating or preventing immune system disorders, e.g. AIDS
 or anemia, cell proliferative disorders, e.g. cancer, or neurological
 disorders.

Claim 5; Page 208; 213pp; English.

ACCS9924 to ACCS9958 encode the human immune response associated proteins
 given in ABR43205 to ABR43239, designated IRAP-1 to IRAP-29 (I). (I) have
 anti-HIV, anti-allergic, anti-inflammatory, antianemic, antiparkinsonian,
 nootropic, anticonvulsant, antiarteriosclerotic, antidiabetic, antipruritic,
 immunosuppressive, antithyroid, cytotoxic, hepatotropic, dermatological,
 antidiabetic, nephrotropic, thyromimetic, neuroprotective, osteopathic,
 antiarthritic, antiparasitic, antihelminthic, antipsoriatic, uropathic,
 ophthalmological, antirheumatic, haemostatic, antibacterial, virucide,
 protozoacide and fungicide activities, and can be used in gene therapy.
 Human IRAP polynucleotides, agonists and antagonists are useful for
 diagnosing, treating or preventing disorders associated with aberrant
 expression of IRAP, such as immune system disorders (e.g. AIDS, asthma,
 allergies, autoimmune thyroiditis, contact dermatitis, Crohn's disease,
 diabetes mellitus, glomerulonephritis, Good pasture's syndrome, gout,
 Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
 multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
 syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis),
 neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or
 epilepsy), developmental disorders (e.g. renal tubular acidosis, anaemia
 or mental retardation), muscle disorder (e.g. cardiomyopathy, myocardiitis
 or dermatomyositis), cell proliferative disorders (e.g. arteriosclerosis,
 atherosclerosis, cirrhosis, hepatitis, paroxysmal nocturnal
 haemoglobinuria, polycythaemia vera, psoriasis, primary
 thrombocytopaenia or cancer), or viral, bacterial, fungal, parasitic,
 protozoan or helminthic infections

Sequence 732 BP; 213 A; 178 C; 178 G; 163 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,28e-10 Length: 732
 Score: 169.50 Matches: 47
 Percent Similarity: 51.61% Conservative: 33
 Best Local Similarity: 30.32% Mismatches: 54

```

Query Match: 23.71% Indels: 21
DB: 7 Gaps: 5
US-10-005-907-2 (1-135) x ACC59951 (1-732)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 134 ATGGGAATTTCTCTGCTGAGAGAAACAGCGCGGAGAGAAACACTCAAGAGATGCTTGG 193
Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30
Db 194 AATGTGAGATGCAAGCCCAACACAGAGAAACATCCAGATGCTGGGATCACCATATCGCT 253
Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 254 GAAGGGTGTCTCTCCCTTCCATGGAATAATACTCAATTTTGAAGAGAGG---CAAGAT 310
Qy 43 GlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySer 62
Db 311 TCCCAACAGAAATGAAGATGTCATCTACTCCATCCAGACATGTTGACCAACC 370
Qy 63 GlySerGluGluValCysTyrThrValIleAsnHis---IleProHisGlnArgSerSer 81
Db 371 TACTCAGAGAGCTGTGTATACCTCATCAATCATCGGTTCTCTGACAAAGGCCATCA 430
Qy 82 LeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGln 101
Db 431 GGGAACTTCTGTAAGAGTACTATGAGATGTTCCCTGCAAGCTGAGACCCAGAGAG 490
Qy 102 PheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCys 121
Db 491 TCCTTGGGAGGAAGTGAAGACTGAGTATTCATCTTCTACATATGCTTCTACAGACCC 550
Qy 122 SerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 551 CATGCCGATCCCAAGAGATGAATATGAACTTCTCATGCTCTAC 595

RESULT 4
ABL99791
ID ABL99791 standard; cDNA; 1643 BP.
XX
AC ABL99791;
XX
DT 03-OCT-2002 (first entry)
XX
DE Human secretory polynucleotide (sptm) 46.
XX
KW Human; ss; gene; secretory protein; secretory polynucleotides; SPTM;
KW SPTM-related disease; somatic gene therapy; germline gene therapy;
KW severe combined immunodeficiency; intracellular parasite protection;
KW fungal parasite; protozoan parasite; cell proliferative disorder; cancer;
KW immune disorder; AIDS; neurological disorder; Parkinson's disease;
KW motor neuron disorder; demyelinating disease; multiple sclerosis;
KW meningitis; abscess; prion diseases; cerebral palsy;
KW neuroskeletal disorder; peripheral nervous system disorder;
KW dermatomyositis; polymyositis; myopathy; myasthenia gravis;
KW mental disorder; Tourette's syndrome.
XX
OS Homo sapiens.
XX
FN WO200220756-A2.
XX
PD 14-MAR-2002.
XX
PF 30-AUG-2001; 2001WO-US027297.
XX
XX 05-SEP-2000; 2000US-0229747P.
PR 05-SEP-2000; 2000US-0229748P.
PR 05-SEP-2000; 2000US-0229749P.
PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.
PR 05-SEP-2000; 2000US-0230016P.
PR 05-SEP-2000; 2000US-0230583P.

Alignment Scores:
Pred. No.: 6,72e-09 Length: 1643
Score: 159.00 Matches: 47
Percent Similarity: 50.64% Conservative: 32
Best Local Similarity: 30.13% Mismatches: 55
Query Match: 22.24% Indels: 22
DB: 6 Gaps: 6
US-10-005-907-2 (1-135) x ABL99791 (1-1643)
Qy 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysPro--- 19
Db 1096 ATGGGAATTTCTCTGCTGAGAGAAACAGCGCGGAGAGAAACACTCAAGAGATGCTTGG 1155
Qy 20 -----LysLysGlyAsnProAspGluGluArgLysArg----- 30

```

(INCY-) INCYTE GENOMICS INC.

Stuart J, Lincoln SE, Altus CM, Dufour GB, Chalup MS, Hillman JL;
Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;
Moriyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;
Gerstin EH, Peralta CH, David MH, Panzer SR, Flores V, Daffo A;
Marwaha R, Chen AJ, Chang SC, Au AP, Imman RR;

WPI; 2002-315658/35.
F-PSDB; ABB97794.

Polynucleotide sequences encoding human secretory proteins useful for gene therapy of e.g. genetic deficiency disorders, cancers, and diseases caused by intracellular parasites.

Claim 1; Page 275-276; 585pp; English.

The invention comprises the amino acid and coding sequences of human secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are useful for treating a disease or condition associated with the expression of functional SPTM. The SPTM DNA sequences are useful for somatic or germline gene therapy to correct a genetic deficiency (e.g. severe combined immunodeficiency). The SPTM DNA sequences are also useful in providing protection against intracellular parasites (e.g. fungal parasites and protozoan parasites). The SPTM DNA and protein sequences are also useful for diagnosing cell proliferative disorders (e.g. cancer, immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's disease), motor neuron disorders, demyelinating diseases (e.g. multiple sclerosis), meningitis, abscesses, prion diseases, cerebral palsy, neuroskeletal disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental disorders (e.g. Tourette's syndrome). cDNA sequences ABL99746 - ABL99929 represent human secretory polynucleotides of the invention

Sequence 1643 BP; 413 A; 382 C; 384 G; 460 T; 0 U; 4 Other;

Db 1156 AATGTGAGATCAAAAGCCCAACAGAGAAATCCAGATCGTGGATCACCATATCGCT 1215
Qy 31 -----GlnGluMetThrThrPheGluArgLysLeuGlnAsp 42
Db 1216 GAAGGGTGTTCCTCCATCGAAATAATACTCATTTTGAAGAAGG---CAAGAT 1272
Qy 43 GlnAspLysLysSerGlnGluValSerSerThrSer---AsnGlnGluAsnGluAsnGly 61
Db 1273 TCCAAACGAAATGAAGATGTCATCTACTCCATCCAGCAGGACAATGTGACCAG 1332
Qy 62 SerGlySerGluGluValCysThrThrValIleAsnHis---IleProHisGlnArgSer 80
Db 1333 ACTTACTCAGAGAGCTGTGTATACCTCATCAATCATCGGTCTCTGTATCAGGCCA 1392
Qy 81 SerLeuSerSerAnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
Db 1393 TCAGGGAACCTGCTGAAGAGTACTATGAGATGTTCCCTGCAAGCTGAGAGACCCAGA 1452
Qy 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
Db 1453 GAGTCCTTGGGAGGAAGTGAAGTCTGAGTATCTCTATATATGCTTCTCTACAGACCCC 1512
Qy 121 CysSerCys---ThrHisGluHisAspTyrGluValValPheProHis 135
Db 1513 AGGCATGCCGATCCCAAGAGATGAATATGAACTTTTTCATGCTCTAC 1560
RESULT 5
ABX05108
ID ABX05108 standard; cDNA; 775 BP.
XX AC ABX05108;
XX DT 17-JAN-2003 (first entry)
XX DE Human novel polynucleotide #123.
XX KW Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
KW neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
KW Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
KW osteoarthritis; periodontal disease; liver fibrosis; viral infection;
KW fungal infection; bacterial infection; autoimmune disease; diabetes;
KW atopic dermatitis.
XX OS Homo sapiens.
XX FN WO200274961-A1.
XX PD 26-SEP-2002.
XX PF 14-MAR-2002; 2002WO-US005109.
XX PR 15-MAR-2001; 2001US-00810173.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang Yt, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QH, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX DR WPI; 2003-040556/03.
XX DR P-PSDB; ABU00030.
XX PT New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX PS Claim 1; SEQ ID NO 123; 235pp; English.
XX CC The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations

CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification but is based on sequence information supplied
CC by the European Patent Office
XX
SQ Sequence 775 BP; 226 A; 202 C; 160 G; 187 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,88e-08 Length: 775
Score: 149.00 Matches: 37
Percent Similarity: 60.19% Conservative: 28
Best Local Similarity: 34.26% Mismatches: 39
Query Match: 20.84% Indels: 4
DB: Gaps: 4
US-10-005-907-2 (1-135) x ABX05108 (1-775)
Qy 31 GlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGluVal 50
Db 329 AAAAAAATACTCATTTTGAAGAAGG---CAAGATTCCTCAAAACGAAATGAAGAATG 385
Qy 51 SerSerThrSer---AsnGlnGluAsnGluArgLysSerGlySerGluGluValCysTyr 69
Db 386 TCATCTACTCCATCCAGCAGGACAATGTGACACAGCTACTCAGAGGAGCTGTGTAT 445
Qy 70 ThrValIleAsnHis---IleProHisGlnArgSerSerLeuSerSerAsnAspAspGly 88
Db 446 ACCCTCATCAATCATCGGTCTCTGTATCAAGCCCATCAGGAGACTCTGCTCAGAGTAC 505
Qy 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108
Db 506 TATGAGATGTTCCCTGCAAACTCAGAGACCCAGAGAGTCTCTGGGAGGAACTGAGACT 565
Qy 109 GlnTyrAlaLeuLeuArgThrSerValSerArgProCysSerCys---ThrHisGluHis 127
Db 566 GAGTATTCACTTCTATCATATGCTTCTACAGACCCCATGCGCATGCCGATCCCCAGAGAT 625
Qy 128 AspTyrGluValValPheProHis 135
Db 626 GAATATGAACCTTCTCATGCTCTAC 649
RESULT 6
AAZ24890
ID AAZ24890 standard; DNA; 1892 BP.
XX AC AAZ24890;
XX DT 02-DEC-1999 (first entry)
XX DE Human secreted protein gene 80 clone HLWAY54.
XX KW Human; secreted protein; fusion protein; gene therapy; protein therapy;
KW diagnosis; tissue; cancer; tumour; neurodegenerative disorder; leukaemia;
KW developmental abnormality; foetal deficiency; blood; allergy; renal; ds;
KW immune system; asthma; lymphocytic disease; brain; hepatic; lymphoma;
KW inflammation; ischaemic shock; Alzheimer's disease; restenosis; AIDS;
KW cognitive disorder; schizophrenia; prostate; obesity; osteoclast; thymus;
KW osteoporosis; arthritis; testis; lung; thyroiditis; thyroid; digestion;
KW endocrine; metabolism; regulation; malabsorption; gastritis; neoplasm.
XX OS Homo sapiens.
XX FN WO9947540-A1.
XX PD 23-SEP-1999.

XX PF 18-MAR-1999; 99WO-US005804.
 XX PR 19-MAR-1998; 98US-0078563P.
 XX PR 19-MAR-1998; 98US-0078566P.
 XX PR 19-MAR-1998; 98US-0078573P.
 XX PR 19-MAR-1998; 98US-0078574P.
 XX PR 19-MAR-1998; 98US-0078576P.
 XX PR 19-MAR-1998; 98US-0078577P.
 XX PR 19-MAR-1998; 98US-0078578P.
 XX PR 19-MAR-1998; 98US-0078579P.
 XX PR 19-MAR-1998; 98US-0078581P.
 XX PR 01-APR-1998; 98US-0080312P.
 XX PR 01-APR-1998; 98US-0080313P.
 XX PR 01-APR-1998; 98US-0080314P.
 XX PA (HUMA-) HUMAN GENOME SCI INC.
 XX PI Ruben SM, Ni J, Rosen CA, Yu G, Young PE, Feng P, Soppet DR;
 PI Wei Y, Endress CA, Duan RD, Kyaw H, Ebner R, Lafleur DW, Olsen HS;
 PI Shi Y, Moore PA;
 XX DR P-PSDB; AA41387.
 XX DR WPI; 1999-562050/47.
 XX PT New isolated human genes, useful for diagnosis and treatment of e.g.
 PT cancers, neurological disorders, immune diseases, inflammation or blood
 PT disorders.
 XX PS Claim 1; Page 346; 484pp; English.
 XX CC This sequence represents a nucleic acid molecule which encodes a secreted
 CC human protein. The gene number, and the clone it is derived from, are
 CC detailed in the descriptor line. The gene can be used to generate fusion
 CC proteins by linking to the gene to a human immunoglobulin FC portion
 CC (e.g. AA24802) for increasing the stability of the fused protein as
 CC compared to the human protein only. The invention relates to 95 novel
 CC genes and their fragments (nucleic acid sequences: AA224811-224907; amino
 CC acid sequences AA41308-Y41404) which are useful for preventing, treating
 CC or ameliorating medical conditions e.g. by protein or gene therapy. Also,
 CC pathological conditions can be diagnosed by determining the amount of the
 CC new polypeptides in a sample or by determining the presence of mutations
 CC in the new polynucleotides. Specific uses are described for each of the
 CC 95 polynucleotides, based on which tissues they are most highly expressed
 CC in (see AA224811 for described uses)
 XX SQ Sequence 1892 BP; 461 A; 565 C; 472 G; 394 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: Length: 1892
 Score: 98.00 Matches: 36
 Percent Similarity: 38.26% Conservative: 21
 Best Local Similarity: 24.16% Mismatches: 62
 Query Match: 13.71% Indels: 30
 DB: 2 Gaps: 4

US-10-005-907-2 (1-135) x AA224890 (1-1892)

QY 5 LeuLeuArgLysLeuSerCysLeuGluAsnGlnLysLysProLysGlyAsnPro 24
 DB 572 CTCTTACATCTCTCTTGTCTCTGAGCCAGGAGCCAGGAGCAGACACAGCAGGAG 631
 QY 25 AspGluGluArgLysGlnGlnMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
 DB 632 CAAGGAGTGAGCAGCAGGAGGAGGCGGACACAAGAACACACAGCAGGAGGGGCGAGAA 691
 QY 45 LysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnLysSerGlySer 64
 DB 692 CAGGAGACGACAGAGAGACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 751
 QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
 DB 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811

QY 81 ---SerLeuSerSerAsnAspGlyTyr----- 89
 DB 812 GAATCTCTATCTCTTAACCTTCTCTTTGCTCCCGGGTACGAGAGTAGAGTCTACT 871
 QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
 DB 872 CCTATGATAATGGAGACATCAGGAGCTCATTCATCAGCCCGGAAATAGATGAATG 931
 QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
 DB 932 AATGAATATATGATGAGAACTCTCTACTGGAGAAACCAAAACCTCGGAGCCTCTGCGAG 991
 QY 115 ThrSerValSerArgProCysSerCys 123
 DB 992 CTGCCCCACACAGAGCCTTGCTGGTGC 1018

RESULT 7
 ADA56130
 ID ADA56130 standard; DNA; 1892 BP.
 XX AC ADA56130;
 XX DT 20-NOV-2003 (first entry)
 XX DE Gene encoding human secreted protein #309.
 XX KW immunosuppressive; antiinflammatory; antiasthmatic; antiallergic;
 KW cytostatic; cerebroprotective; neuroprotective; nootropic;
 KW cardiovascular; antiarteriosclerotic; gene therapy;
 KW human secreted protein; immune disorder; inflammation;
 KW respiratory bowel disease; cancer; CNS disorder; neurodegenerative disorders;
 KW inflammatory disorder; nephritis; Crohn's disease; asthma; allergy;
 KW multiple sclerosis; ischaemic brain injury; Parkinson's disease;
 KW Alzheimer's disease; atherosclerosis; myocarditis; chromosome mapping;
 KW triple helix formation; antisense gene therapy; forensic biology; ds;
 KW gene.
 XX OS Homo sapiens.
 XX PN WO2002102994-A2.
 XX PD 27-DEC-2002.
 XX PF 19-MAR-2002; 2002WO-US008278.
 XX PR 21-MAR-2001; 2001US-0277340P.
 XX PR 19-JUL-2001; 2001US-0306171P.
 XX PR 13-NOV-2001; 2001US-0331287P.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX PA Rosen CA, Ruben SM;
 XX PI WPI; 2003-167512/16.
 XX DR P-PSDB; ADA57026.
 XX FT New human secreted polypeptides and polynucleotides, useful for
 PT diagnosing, treating or preventing e.g. immune disorders, inflammatory
 PT conditions, respiratory disorders, cancers, CNS disorders, or
 PT neurodegenerative disorders.
 XX PS Claim 21; SEQ ID NO 319; 1754pp; English.
 XX CC The invention relates to 592 new human secreted polypeptides useful for
 CC diagnosing, treating or preventing e.g. immune disorders, inflammatory
 CC conditions, respiratory disorders, cancers, CNS disorders, or
 CC neurodegenerative disorders, or polypeptides comprising an amino acid
 CC sequence at least 95% identical to the new sequences. The polypeptides,
 CC antibodies or antibody fragments that bind to the polypeptides, nucleic
 CC acids encoding the polypeptides, agonists or antagonists that binds to
 CC the polypeptide, are useful in preparing diagnostic or pharmaceutical
 CC compositions for diagnosing, treating or preventing an e.g. immune

disorders, inflammatory conditions (e.g. inflammatory bowel disease, nephritis or Crohn's disease), respiratory disorders (e.g. asthma and allergy), cancers (e.g. gastric, ovarian or lung cancer), CNS disorders (e.g. multiple sclerosis or ischaemic brain injury), neurodegenerative disorders (e.g. Parkinson's disease or Alzheimer's disease), and cardiovascular disorders (e.g. atherosclerosis or myocarditis). The polynucleotides are useful for chromosome identification, chromosome mapping, for controlling gene expression through triple helix formation or antisense DNA or RNA, in gene therapy, for identifying individuals from minute biological samples, in forensic biology, and as hybridization probes. The polypeptides are useful for as molecular weight markers on sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) gels, to raise antibodies, for testing biological activities, and for treating or preventing neural disorders, immune system disorders, muscular, reproductive, gastrointestinal, pulmonary, cardiovascular, renal, proliferative and/or cancerous diseases. This sequence corresponds to a gene encoding one of the polypeptide of the invention. Note: The sequence data for this patent did form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1892 BP; 461 A; 565 C; 472 G; 394 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 0.156 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservative: 21
Best Local Similarity: 24.16% Mismatches: 50
Query Match: 13.71% Indels: 32
DB: 7 Gaps: 4

US-10-005-907-2 (1-135) x ADA56130 (1-1892)

QY 5 LeuLeuArgLysLeuSerCysLeuGluAsnGlnLysLysProLysLysGlyAsnPro 24
Db 572 CTCTCAATCAATCTCTTGTCTCCCTGGGAAGCCAGGAGCCGACGACAGCAGCAGGAG 631
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 632 CAAGGAGTGAGCAGCAGGAGGAGCCGACACAGACAGCAGCAGGAGGAGGAGGAG 691
QY 45 LysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlnGlySerGlySer 64
Db 692 CAGGAGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 751
QY 65 GluGlu-----ValCysThrThrValIleAsnHisIleProHisGlnArgSer 80
Db 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
QY 81 ---SerLeuSerSerAsnAspGlyTyr----- 89
Db 812 GAATCTCTATCTTCTTACCTCTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
Db 872 CCTATGATATGAGAACATCCAGGAGCTCATTCGATCAGCCAGGAGGAGGAGGAGGAG 931
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 932 AATGAATATATGATGAGAACTCTACTGGAGAAACCAAAACCTGGCAGCTCTCTGAG 991
QY 115 ThrSerValSerArgProCysSerCys 123
Db 992 CTGCCCCACAGCAGCTTGTGTGTG 1018

RESULT 8

ADA39940

ID ADA39940 standard; cDNA; 1892 BP.

XX

AC ADA39940;

XX 20-NOV-2003 (first entry)

DT

XX

DE Human secreted protein encoding cDNA.

XX Human; secreted protein; cancer; hyperproliferative disorder;
KW rheumatoid arthritis; autoimmune disorder; haematopoietic disorder;
KW anemia; allergic reaction; asthma; cardiovascular disorder;
KW wound healing; cytostatic; immunosuppressive; neurotropic; neuroprotective;
KW antiviral; antiallergic; hepatotropic; antidiabetic; antiinflammatory;
KW vulnery; candiant; gene therapy; ss.

OS Homo sapiens.

XX WO2002102993-A2.

PN 27-DEC-2002.

XX 19-MAR-2002; 2002WO-US008123.

XX 21-MAR-2001; 2001US-0277340P.

XX 19-JUL-2001; 2001US-0306171P.

XX 13-NOV-2001; 2001US-0331287P.

XX (HUMA-) HUMAN GENOME SCI INC.

XX Rosen CA, Ruben SM;

XX WPI; 2003-175238/17.

XX New human secreted proteins and nucleic acid molecules, useful for preparing a diagnostic or pharmaceutical composition for diagnosing, preventing or treating cancer or other hyperproliferative disorder, asthma, allergies or AIDS.
PS Claim 9; SEQ ID NO 322; 3205pp; English.

XX The invention relates to novel genes ADA39629-ADA40555 and proteins ADA40566-ADA41501 for human secreted proteins, useful for preventing, treating or ameliorating medical conditions e.g. by protein or gene therapy. The polypeptides, nucleic acid molecules, antibodies or their fragments, and agonists or antagonists that bind to the polypeptide are useful for preparing a diagnostic or pharmaceutical composition for diagnosing or treating cancer or other hyperproliferative disorder. The polypeptides and nucleic acid molecules are also useful for detecting, preventing, diagnosing, prognosticating, treating or ameliorating cancer or other hyperproliferative disorders including neoplasms, autoimmune disorders (e.g. diabetes, rheumatoid arthritis, systemic lupus erythematosus, multiple sclerosis, autoimmune thyroiditis or haemolytic anaemia), haematopoietic or haematological disorders (e.g. anaemia, thrombocytopenia), allergic reactions including asthma or eczema, inflammatory disorders (e.g. ischaemia-reperfusion injury, inflammatory bowel disease or Crohn's disease), neurodegenerative disorders (e.g. Alzheimer's disease or Parkinson's disease), cardiovascular disorders (e.g. atherosclerosis, myocarditis), infectious diseases (bacterial, fungal or viral infections including HIV/AIDS), or wound healing and disorders of epithelial cell proliferation. The nucleic acids are also useful for chromosome identification, radiation hybrid mapping or long-range restriction mapping, as molecular weight markers, or as hybridization or diagnostic probes. The polypeptides and antibodies are useful for providing immunological probes for differential identification of the tissues immunohistochemistry assays. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 1892 BP; 461 A; 565 C; 472 G; 394 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 0.156 Length: 1892

Score: 98.00 Matches: 36

Percent Similarity: 38.26% Conservative: 21

Best Local Similarity: 24.16% Mismatches: 50

Query Match: 13.71% Indels: 32

DB: 7 Gaps: 4

Alignment Scores:

Pred. No.: 0.587 Length: 1824
 Score: 93.00 Matches: 36
 Percent Similarity: 37.20% Conservatives: 25
 Best Local Similarity: 21.95% Mismatches: 51
 Query Match: 13.01% Indels: 52
 DB: 4 Gaps: 5
 US-10-005-907-2 (1-135) x ABL27495 (1-1824)
 QY 3 AsnTyrLeuLeuArgLysLeuSerCys---LeuGlyGluAenGlnLysLysProLysLys 21
 Db 940 AACGAGGATTGCGGATCTGTGTCTCTGACGACGACCGCGGAGAGGACGCAAG 999
 QY 22 -----GlyAsnProAspGluGluArgLysArgGlnGluMet 33
 Db 1000 CTGGCGGGGAATGCAAGAGATTGGCGGATACACGCCCGCGTGTGCGGAGGAGTG 1059
 QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluVal----- 50
 Db 1060 GCCCGGTACCAAGAACTACGTGAGTGTATGATGATGATGATGATGATGATGATGAT 1119
 QY 50 ----- 50
 Db 1120 AATCTAGAGTCAGGAGCTGCTGCTTACTTGACGAGGAGCGTGCCTATGGCGGCA 1179
 QY 51 -----SerSerThrSerAenGlnGluAenGluAenGlySer 62
 Db 1180 AATCGCTGTGCGCGGTGGGAGGACGACCTAGGAATGCGTTCGAGATGACGCGCAT 1239
 QY 63 GlySerGlu-----GluValCysTyrThrValIleAenHisIleProHis 77
 Db 1240 GGTTCAGCTCCACGACGATGAGATGAGACCATCCCGCTTGAGGAACATATGCGAG 1299
 QY 78 GlnArgSer-----SerLeuSerSerAenAspAspGlyTyrGlu 90
 Db 1300 CAGCGGCAACTACCCAGGATCTACGTCTGCGCCACGCTGAGACGACGCTTCAG 1359
 QY 91 AsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThrGluTyr 110
 Db 1360 TATGTGGCTCAGTGGAGCGCGCATCCAGCACTGGAGGAGGAGCGCACCGCCACG 1419
 QY 111 AlaLeuLeuArg 114
 Db 1420 GCCCATCTACAA 1431

RESULT 15

AAS71153
 ID AAS71153 standard; cDNA; 654 BP.
 AC AAS71153;
 XX
 DT 13-FEB-2002 (first entry)
 XX
 DE DNA encoding novel human diagnostic protein #6957.
 XX
 KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
 XX food supplement; medical imaging; diagnostic; genetic disorder; ss.
 XX
 OS Homo sapiens.
 XX
 PN W0200175067-A2.
 XX
 PD 11-OCT-2001.
 XX
 PF 30-MAR-2001; 2001WO-US0008631.
 XX
 PR 31-MAR-2000; 2000US-00540217.
 PR 23-AUG-2000; 2000US-00649167.
 XX
 PA (HYSE-) HYSEQ INC.
 XX
 PI Drmanac RT, Liu C, Tang YT;

XX

DR WPI; 2001-639362/73.
 DR P-PSDB; ABG06966.
 XX

PT New isolated polynucleotide and encoded polypeptides, useful in
 PT diagnostics, forensics, gene mapping, identification of mutations
 PT responsible for genetic disorders or other traits and to assess
 PT biodiversity.
 XX

PS Claim 1; SEQ ID NO 6957; 103pp; English.
 XX

CC The invention relates to isolated polynucleotide (I) and polypeptide (II)
 CC sequences. (I) is useful as hybridisation probes, polymerase chain
 CC reaction (PCR) primers, oligomers, and for chromosome and gene mapping,
 CC and in recombinant production of (II). The polynucleotides are also used
 CC in diagnostics as expressed sequence tags for identifying expressed
 CC genes. (I) is useful in gene therapy techniques to restore normal
 CC activity of (II) or to treat disease states involving (II). (II) is
 CC useful for generating antibodies against it, detecting or quantitating a
 CC polypeptide in tissue, as molecular weight markers and as a food
 CC supplement. (II) and its binding partners are useful in medical imaging
 CC of sites expressing (II). (I) and (II) are useful for treating disorders
 CC involving aberrant protein expression or biological activity. The
 CC polypeptide and polynucleotide sequences have applications in
 CC diagnostics, forensics, gene mapping, identification of mutations
 CC responsible for genetic disorders or other traits to assess biodiversity
 CC and to produce other types of data and products dependent on DNA and
 CC amino acid sequences. AAS64197-AAS94564 represent novel human diagnostic
 CC coding sequences of the invention. Note: The sequence data for this
 CC patent did not appear in the printed specification, but was obtained in
 CC electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 654 BP; 304 A; 81 C; 181 G; 88 T; 0 U; 0 Other;
 XX

Alignment Scores:

Pred. No.: 0.198 Length: 654
 Score: 92.00 Matches: 35
 Percent Similarity: 41.01% Conservatives: 22
 Best Local Similarity: 25.18% Mismatches: 53
 Query Match: 12.87% Indels: 30
 DB: 5 Gaps: 3

US-10-005-907-2 (1-135) x AAS71153 (1-654)

QY 14 GluAenGlnLysLysProLysGlyGlyAsnProAspGluGluArgLysArgGlnGluMet 33
 Db 169 GAAGGAA 228
 QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerSerThr 53
 Db 229 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 288
 QY 54 SerAenGlnGluAenGluAenGlySerGlySerGluGluValCysTyrThrValIleAen 73
 Db 289 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 330
 QY 74 HisIleProHisGlnArgSerSerLeuSer----- 83
 Db 331 GAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGAAGA 390
 QY 84 SerAenAspAspGlyTyrGluAenIleAspSerLeuThrArg----- 97
 Db 391 TCAGACATCAACGGATTGTGAAA-ATTACTTCCAGTACCCCTATTATTAGCAGTACTAGA 449
 QY 98 -----LysValArgGlnPheArgGluArgSerGluThrGluTyr 110
 Db 450 GGATGCGTTTACGAGTCTCTGGAAAACACGACATAAAGGCAAGAAGAAGAAGAAGAAG 509
 QY 111 AlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGluHisAspTyr 129
 Db 510 GAAACAGGGGATTCAGATGATCGAGAGCAAAAGGATGTTCCACAGATGATGACTAT 566

Search completed: April 22, 2004, 17:20:13
Job time : 360 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: April 22, 2004, 17:57:48 ; Search time 334 Seconds
(without alignments)
1822.349 Million cell updates/sec

Title: US-10-005-907-2

Perfect score: 715

Sequence: 1 MGNYLKLSCLGENQKPK.....SVSRPSCSTHEHDYEVFFH 135

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO.spool_p/US10005907/runat_21042004_113756_11532/app_query.fasta_1.327
-DB=Published Applications NA -Qfmt=fastap -SURFIX=rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US10005907@cgn1.1.333 @runat_21042004_113756_11532
-NCPU=6 -ICPU=1 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:

1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/US07_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:
12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:
15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:
16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:
17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
------------	-------	--------------------	----	-------------

1	715	100.0	3762	15	US-10-005-907-1	Sequence 1, Appli
2	98	13.7	1892	13	US-10-653-595-90	Sequence 90, Appl
3	98	13.7	1892	13	US-03-397-945-90	Sequence 30, Appl
4	96	13.4	2301	15	US-10-242-943-3	Sequence 3, Appli
5	93	13.0	1903	13	US-10-424-599-8569	Sequence 8569, Ap
6	92	12.9	6324	16	US-10-115-831-54	Sequence 54, Appl
7	91.5	12.8	2862	13	US-10-263-929-22	Sequence 22, Appl
8	91.5	12.8	5109	13	US-10-342-887-1552	Sequence 1552, Ap
9	91.5	12.8	5109	13	US-10-172-118-1552	Sequence 1552, Ap
10	91.5	12.8	5261	10	US-09-814-353-19577	Sequence 19577, A
11	91.5	12.8	7488	15	US-10-311-455-1085	Sequence 1085, Ap
12	90.5	12.7	651	13	US-10-261-175A-9	Sequence 9, Appl
13	90	12.6	1173	13	US-10-027-632-10372	Sequence 10372, A
14	90	12.6	1173	16	US-10-027-632-10372	Sequence 10372, A
15	90	12.6	1671	13	US-10-296-115-693	Sequence 693, App
16	90	12.6	1899	10	US-09-746-783-77	Sequence 77, Appl
17	90	12.6	7992	10	US-09-893-519A-140	Sequence 140, App
18	90	12.6	8491	15	US-10-133-013-260	Sequence 260, App
19	90	12.6	8493	14	US-10-071-766-51	Sequence 51, Appl
20	90	12.6	8503	15	US-10-037-270-130	Sequence 130, App
21	90	12.6	8503	16	US-10-117-722-130	Sequence 130, App
22	89.5	12.5	1071	13	US-09-933-767-118	Sequence 118, App
23	89.5	12.5	1071	13	US-10-004-860-118	Sequence 118, App
24	89.5	12.5	1071	15	US-10-023-282-118	Sequence 118, App
25	89.5	12.5	1107	13	US-10-424-599-112896	Sequence 112896,
26	89.5	12.5	9415	15	US-10-311-455-268	Sequence 268, App
27	89	12.4	864	16	US-10-320-800-49	Sequence 49, Appl
28	88	12.3	99934	16	US-10-085-117-73	Sequence 73, Appl
29	87	12.2	3720	14	US-10-007-805-576	Sequence 576, App
30	87	12.2	3720	15	US-10-076-622-576	Sequence 576, App
31	87	12.2	3720	15	US-10-124-805-576	Sequence 576, App
32	87	12.2	123192	15	US-10-175-523-71	Sequence 71, Appl
33	86.5	12.1	776	13	US-10-424-599-23224	Sequence 23224, A
34	86.5	12.1	1181	13	US-10-415-134-3	Sequence 3, Appli
35	86.5	12.1	1302	9	US-09-879-957-17	Sequence 17, Appl
36	86.5	12.1	7095	13	US-10-205-331-67	Sequence 67, Appl
37	86	12.0	1632	16	US-10-085-117-54	Sequence 54, Appl
38	86	12.0	1725	15	US-10-198-846-10857	Sequence 10857, A
39	86	12.0	1886	15	US-10-262-666-41	Sequence 41, Appl
40	86	12.0	1886	16	US-10-085-117-53	Sequence 53, Appl
41	86	12.0	3883	15	US-10-177-293-331	Sequence 331, App
42	85.5	12.0	1482	9	US-09-938-842A-585	Sequence 585, App
43	85.5	12.0	1482	11	US-09-938-842A-585	Sequence 141192,
44	85.5	12.0	2257	13	US-10-424-599-141192	Sequence 141192,
45	85.5	12.0	2264	13	US-10-424-599-13803	Sequence 13803, A

ALIGNMENTS

RESULT 1
US-10-005-907-1
; Sequence 1, Application US/10005907
; Publication No. US2003016881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US2003016881A1A, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAINT CRI
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005.907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 1
; LENGTH: 3762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(432)

```

; OTHER INFORMATION:
US-10-005-907-1
Alignment Scores:
Pred. No.: 2,31e-85 Length: 3762
Score: 715.00 Matches: 135
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 15 Gaps: 0

US-10-005-907-2 (1-135) x US-10-005-907-1 (1-3762)
QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGluAAsnGlnLysLysProLys 20
DB 25 ATGGGAATTAATCTCTCGGAAATCTAGTTGCTGGGAGAGAAATCAAAAGAAAGCCCAAG 84
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 85 AAAGGAACCCAGATGAGGAAGAAAGAAAGCGGAGGAATGACTACATTTGAAGAAACTT 144
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnLysAsnGluAsn 60
DB 145 CAAGATCAAGATAAGAAAGCCAGAAAGTTTCATCCACTTCTAATCAGGAAAGCGAAT 204
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 205 GGCATGGTTCTGAAGAGTGTGTACTGCTATTATCATCATCCCTCCATCAGAGATCC 264
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
DB 265 TCCCTGAGCTCCAATGATGCTGCTATGAGAACATGACTCCCTCACAAGGAAAGTGAGA 324
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArgPro 120
DB 325 CAGTTTAGAAGAAAGTCTAGACAGAAATATGCCCTCTTAGACTTCTGTAGAGCCT 384
QY 121 CysSerCysThrHisGluHisAspTyrGluValValPheProHis 135
DB 385 TGTTCTGCACCCATGAGCATGATTATGAAGTTGTGTTCACAC 429

RESULT 2
US-10-653-595-90
; Sequence 90, Application US/10653595
; Publication No. US20040048304A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et. al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027P1C1
; CURRENT APPLICATION NUMBER: US/10/653,595
; PRIOR FILING DATE: 2003-09-03
; PRIOR APPLICATION NUMBER: US 09/397945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080,312
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/078,578
; PRIOR FILING DATE: 1998-03-19
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470

; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 90
; LENGTH: 1892
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-653-595-90
Alignment Scores:
Pred. No.: 0.0109 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservative: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 13 Gaps: 4

US-10-005-907-2 (1-135) x US-10-653-595-90 (1-1892)
QY 5 LeuLeuArgLysLeuSerCysLeuGluAAsnGlnLysLysProLysLysGlyAsnPro 24
DB 572 CTCTACATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 631
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
DB 632 CAAGGAGTGGAGCACACAGGAGGAGCGGACACACAGAAACACACAGCAGCAGAGGGGAGAA 691
QY 45 LysLysSerGlnGluValSerThrSerAsnGlnGluAsnGlnAsnGlySerGlySer 64
DB 692 CAGGAAGACAAAGAGAGAGAAACAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 751
QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
DB 752 AAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 811
QY 81 ---SerLeuSerSerAsnAspAspGlyTyr----- 89
DB 812 GAATCTCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 871
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
DB 872 CCTATGATTAATGAGAACATCCAGGAGCTCATTCGATCAGCCCGAGGAAATAGATGAATG 931
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
DB 932 AATGAATATATGATGAGAACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 991
QY 115 ThrSerValSerArgProCysSerCys 123
DB 992 CTGCCCCACACAGAGCCCTTGCTGTGTC 1018

RESULT 3
US-09-397-945-90
; Sequence 90, Application US/09397945
; Publication No. US20030065139A1
; GENERAL INFORMATION:
; APPLICANT: Human Genome Sciences, Inc. et al.
; TITLE OF INVENTION: 95 Human secreted proteins
; FILE REFERENCE: P2027P1
; CURRENT APPLICATION NUMBER: US/09/397,945
; PRIOR FILING DATE: 1999-09-17
; PRIOR APPLICATION NUMBER: PCT/US99/05804
; PRIOR FILING DATE: 1999-03-18
; PRIOR APPLICATION NUMBER: 60/078,566
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,576
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,573
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,574
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/078,579
; PRIOR FILING DATE: 1998-03-19
; PRIOR APPLICATION NUMBER: 60/080,314
; PRIOR FILING DATE: 1998-04-01
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470

```

;; PRIOR APPLICATION NUMBER: 60/080,312
;; PRIOR FILING DATE: 1998-04-01
;; PRIOR APPLICATION NUMBER: 60/078,578
;; PRIOR FILING DATE: 1998-03-19
;; PRIOR APPLICATION NUMBER: 60/078,581
;; PRIOR FILING DATE: 1998-03-19
;; PRIOR APPLICATION NUMBER: 60/078,577
;; PRIOR FILING DATE: 1998-03-19
;; PRIOR APPLICATION NUMBER: 60/078,563
;; PRIOR FILING DATE: 1998-03-19
;; PRIOR APPLICATION NUMBER: 60/080,313
;; PRIOR FILING DATE: 1998-04-01
;; NUMBER OF SEQ ID NOS: 470
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 90
;; LENGTH: 1892
;; TYPE: DNA
;; ORGANISM: Homo sapiens
US-09-397-945-90

Alignment Scores:
Pred. No.: 0.0109 Length: 1892
Score: 98.00 Matches: 36
Percent Similarity: 38.26% Conservative: 21
Best Local Similarity: 24.16% Mismatches: 62
Query Match: 13.71% Indels: 30
DB: 13 Gaps: 4

US-10-005-907-2 (1-135) x US-09-397-945-90 (1-1892)

QY 5 LeuLeuArgLysLeuSerCysLeuGluAenGlnLysLysProLysLysGlyAsnPro 24
Db 572 CTCCTACATCTCTCTGAGCCAGGAGCAAGCGCCAGACCAAGCAGGAG 631
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 632 CAAGGAGTGGAGCAGCAGGAGCCGACACAAAGACCAAGCAGGAGGGGCAGAAA 691
QY 45 LysLysSerGlnGluValSerSerThrSerAsnGlnGluAenGluAenGlySerGlySer 64
Db 692 CAGGAAGACGACAGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 751
QY 65 GluGlu-----ValCysThrThrValIleAsnHisIleProHisGlnArgSer 80
Db 752 AAGGAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGT 811
QY 81 ---SerLeuSerSerAsnAspAspGlyTyr----- 89
Db 812 GAATCTCTATCTCTTAACTCTTCTGCTCCCGGTACGAGAGTAGAGTCTACT 871
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
Db 872 CCTATGATAATGAGAGACATCCAGGAGCTCATTCGATCAGCCAGGAAATAGATGAATG 931
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 932 AATGAATATATATGATGAGAACTCTACTCGAGAAACCAACCCCTGGCAGCTCTCGAG 991
QY 115 ThrSerValSerArgProCysSerCys 123
Db 992 CTGCCCCACACAGAGCTTGTCTGCTGTC 1018

RESULT 4

US-10-242-943-3
; Sequence 3, Application US/10242943
; Publication No. US20030087412A1
; GENERAL INFORMATION:
; APPLICANT: Nandabalan, Krishnan
; APPLICANT: Yang, Meija
; APPLICANT: Schulz, Vincent P.
; TITLE OF INVENTION: NIK1 PROTEIN AND NIK1 PROTEIN COMPLEXES
; FILE REFERENCE: 15966-521 NIK1 protein complexes
; CURRENT APPLICATION NUMBER: US/10/242,943

;; CURRENT FILING DATE: 2002-09-13
;; PRIOR APPLICATION NUMBER: US/09/167,206
;; PRIOR FILING DATE: 1998-10-06
;; NUMBER OF SEQ ID NOS: 26
;; SOFTWARE: PatentIn Ver. 2.0
;; SEQ ID NO 3
;; LENGTH: 2301
;; TYPE: DNA
;; ORGANISM: Homo sapiens
;; FEATURE:
;; NAME/KEY: CDS
;; LOCATION: (233)..(2155)
US-10-242-943-3

Alignment Scores:

Pred. No.: 0.0268 Length: 2301
Score: 96.00 Matches: 29
Percent Similarity: 41.41% Conservative: 24
Best Local Similarity: 22.66% Mismatches: 51
Query Match: 13.43% Indels: 24
DB: 15 Gaps: 3

US-10-005-907-2 (1-135) x US-10-242-943-3 (1-2301)

QY 14 GluAenGlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
Db 17 GAGGAGCAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 76
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerThr 53
Db 77 GAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
QY 54 SerAsnGlnGluAenGlnGlySerGlySerGluGluValCys----- 68
Db 137 GAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 196
QY 69 -----TyrThrValIleAsn 73
Db 197 AGGAACCGGAGCCGAGCAGTAGCTGCTGGTGGCCACCATCGGATCACCACCATCGAG 256
QY 74 HisIleProHisGlnArgSerSerLeuSerSerAsn---AspAspGlyTyrGluAenIle 92
Db 257 GCGGTGAAGCGCAAGATCCAGGTTCTTCAGCAGCAGGAGGAGGAGGAGGAGGAGGAGT 316
QY 93 AspSerLeuThrArgLysVal-----ArgGlnPheArgGluArgSerGluThrGlu 109
Db 317 GAGCGCTCCAGGAGGAAGTTGAGGAGAAAGCGGGCCCGGGAACAGGCTGAGGCTGAG 376
QY 110 TyrAlaLeuLeuArgThrSerVal 117
Db 377 GTGGCTCTCTTGAACCGTAGGATC 400

RESULT 5

US-10-424-599-8569
; Sequence 8569, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 8569
; LENGTH: 1903
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_107747C.1

US-10-424-599-8569

Alignment Scores:
Pred. No.: 0.0518
Score: 93.00
Percent Similarity: 43.33%
Best Local Similarity: 30.00%
Query Match: 13.01%
DB: 13

Length: 1903
Matches: 36
Conservative: 16
Mismatch: 52
Indels: 16
Gaps: 5

US-10-005-907-2 (1-135) x US-10-424-599-8569 (1-1903)

QY 3 AsnTyrLeuLeuArgLysLeuSerCysLeu---GlyGluAsnGln----- 16
DB 710 AACAGTCTGATTGAGAGATGACACTACTGAGAGGAGAAACGAACTGGTAAATGAAG 769
QY 17 -----LysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
DB 770 TTTCTGAACAGAAACATCAATAGGCTCAATCAGATTGAGAGTATCAGTCAAGTGAAGGA 829
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerThr 53
DB 830 GAAATGAGAGAGAGAGAGAGTGGTATGAGGACAAATAATACAGGAAAGATCACCCAGT 889
QY 54 SerAsnGlnGluAsnGluAsnGlySerGly---SerGluGluValCysTyrThrValile 72
DB 890 CACCAACAGACACAAAGAAATGGCAAGACTACAAAGAGATGCTGTGACAGAAATAT 949
QY 73 AsnHisIleProHisGlnArgSerSerLeuSerSerAsn---AspAspGlyTyrGluAsn 91
DB 950 AACAGTCCAGTCCAGAAAGAGTGTCCAAAGAGGTCAATTGGATTGAGGGTATGAATCA 1009
QY 92 IleAspSerLeuThrArgLysValArgGlnPhe-----ArgGluArgSer 106
DB 1010 AGTGAAGAGAAAGATCA 1069

RESULT 6

Sequence 54, Application US/10115831
Publication No. US20030219743A1
GENERAL INFORMATION:
APPLICANT: Tang, Y. Tom
APPLICANT: Liu, Chenghua
APPLICANT: Asundi, Vinod
APPLICANT: Ren, Feiyao
APPLICANT: Drmanac, Radoje T.
TITLE OF INVENTION: Polypeptides
FILE REFERENCE: 792CIP2ADIV
CURRENT APPLICATION NUMBER: US/10/115,831
CURRENT FILING DATE: 2002-04-02
PRIOR APPLICATION NUMBER: 09/667,298
PRIOR FILING DATE: 2000-09-22
PRIOR APPLICATION NUMBER: 09/577,408
PRIOR FILING DATE: 2000-05-18
NUMBER OF SEQ ID NOS: 178
SOFTWARE: pt_fl_genes Version 2.0
SEQ ID NO 54
LENGTH: 6324
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (1)..(3573)
US-10-115-831-54

Alignment Scores:
Pred. No.: 0.392
Score: 92.00
Percent Similarity: 38.60%
Best Local Similarity: 23.68%
Query Match: 12.87%
DB: 16

Length: 6324
Matches: 27
Conservative: 17
Mismatch: 50
Indels: 20
Gaps: 3

US-10-005-907-2 (1-135) x US-10-115-831-54 (1-6324)

QY 14 GluAsnGlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
DB 337 GACAATGACGAGCCAGGCGGATGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 396
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerThr 53
DB 397 GAGGAGATCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 456
QY 54 SerAsnGlnGluAsnGluAsnGlySerGlySerGluGluVal----- 67
DB 457 GAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 516
QY 68 -----CysTyrThrValileAsnHisIleProHisGlnArgSerSerLeuSerSerAsn 85
DB 517 ATGAATTGTCAAT-----ACTCGAATATGCAAGACACAGAAAGGATGATAACAAT 570
QY 86 AspAspGlyTyrGluAsnIleAspSerLeuThrArgLysVal----- 99
DB 571 AATGACGAATATGACAAATACGATGAATGCTGGCCAAAGTCAATTTTAAACCTCGGCAA 630
QY 100 -----ArgGlnPheArgGluArgSerGluThrGlu 109
DB 631 ATCGCTGAGGATGACGCTACCGGCCAGGACTGAGTCAGAA 672

RESULT 7

US-10-263-929-22
Sequence 22, Application US/10263929
Publication No. US2004006733A1
GENERAL INFORMATION:
APPLICANT: Kim, Jaeseob
APPLICANT: Galant, Ron
TITLE OF INVENTION: Alzheimer's Disease Linked Genes
FILE REFERENCE: LSD-07417
CURRENT APPLICATION NUMBER: US/10/263,929
CURRENT FILING DATE: 2002-10-03
NUMBER OF SEQ ID NOS: 213
SOFTWARE: PatentIn version 3.2
SEQ ID NO 22
LENGTH: 2862
TYPE: DNA
ORGANISM: Homo sapiens
US-10-263-929-22

Alignment Scores:
Pred. No.: 0.148
Score: 91.50
Percent Similarity: 47.90%
Best Local Similarity: 26.89%
Query Match: 12.80%
DB: 13

Length: 2862
Matches: 32
Conservative: 25
Mismatch: 45
Indels: 17
Gaps: 4

US-10-005-907-2 (1-135) x US-10-263-929-22 (1-2862)

QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
DB 394 ATCGGCCAGTCGTTGTTGAAGAGAGAAAGACCTTAACGAGAGAGAAACGAGCTGCTGGAG 453
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu 40
DB 454 GAG-----CAGGTGGAAACACATCAGGAGGAGGTGTCTCAGCTCCGGCATGAGCTG 504
QY 41 GlnAspGlnAspLysSerGlnGluValSerThrSerAsnGlnGluAsnGluAsn 60
DB 505 TCCATGAAGATGAGCTGCTTCACTTACACAGCGCTCGGAGGAG----- 552
QY 61 GlySerGlySerGluGluValCysTyrThrThrValileAsnHisIleProHisGlnArgSer 80
DB 553 ---AGTGAGCCCGAGTCCGTTTCTCAACC-----CCGTTGAAGAGGAAT 594
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100

US-10-172-118-1552

```
; CURRENT APPLICATION NUMBER: US/09/814,353
; CURRENT FILING DATE: 2001-03-21
; PRIOR APPLICATION NUMBER: US 60/191,031
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: US 60/207,124
; PRIOR FILING DATE: 2000-05-25
; PRIOR APPLICATION NUMBER: US 60/211,940
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: US 60/216,820
; PRIOR FILING DATE: 2000-07-07
; PRIOR APPLICATION NUMBER: US 60/220,661
; PRIOR FILING DATE: 2000-07-25
; PRIOR APPLICATION NUMBER: US 60/257,672
; PRIOR FILING DATE: 2000-12-21
; NUMBER OF SEQ ID NOS: 22037
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 19577
; LENGTH: 5261
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: 1, 2, 3, 4, 5, 6, 5250, 5251, 5252, 5253, 5254, 5255, 5256,
; LOCATION: 5257, 5258, 5259, 5260, 5261
; OTHER INFORMATION: n = A,T,C or G
US-09-814-353-19577

Alignment Scores:
Pred. No.: 0.352 Length: 5261
Score: 21.50 Matches: 32
Percent Similarity: 47.90% Conservative: 25
Best Local Similarity: 26.89% Mismatches: 45
Query Match: 12.80% Indels: 17
DB: 10 Gaps: 4

US-10-005-907-2 (1-135) x US-09-814-353-19577 (1-5261)
QY 1 MetGlyAsnTyrLeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLys 20
Db 610 ATCGCCAGTCGTTGTTGAAGAACACCAAGACCCCTAACCCGAGAGGAGCTGCTCGAG 669
QY 21 LysGlyAsnProAspGluGluArgLysArgGlnGluMetThrPheGluArgLysLeu 40
Db 670 GAG-----CAGGTGGACACATCAGGAGGAGGTGCTCAGTCCGGCATGAGCTG 720
QY 41 GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGluAsnGluAsn 60
Db 721 TCCATGAAGAGATGAGCTCTCAGTCTTACACCAAGCGCTCGGAGGAG-----768
QY 61 GlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 769 ---AGTGAGCCGAGTCGTTGCTCAACC-----CGTTCAAGAGGAAT 810
QY 81 SerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
Db 811 GAGTCGCTCTCTCAGTCCAGAAATTAATTTGATTTGATTTCTTTCAAAAGAGCTGAAA 870
QY 101 GlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArgThrSerValSerArg 119
Db 871 GACCTTGAAGAG-----GAGATGTTGATTTCTGATTCGAGGCGCCAGCCAG 915

RESULT 11
US-10-311-455-1085/c
; Sequence 1085, Application US/10311455
; Publication No. US20030143606A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPNBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determ
; FILE REFERENCE: 5013.1014
; CURRENT APPLICATION NUMBER: US/10/311,455
```

```
; CURRENT FILING DATE: 2002-12-16
; PRIOR APPLICATION NUMBER: PCI/EP01/07537
; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 1085
; LENGTH: 7488
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-1085

Alignment Scores:
Pred. No.: 0.593 Length: 7488
Score: 91.50 Matches: 29
Percent Similarity: 44.22% Conservative: 36
Best Local Similarity: 19.73% Mismatches: 51
Query Match: 12.80% Indels: 31
DB: 15 Gaps: 5

US-10-005-907-2 (1-135) x US-10-311-455-1085 (1-7488)
QY 14 GluAsnGlnLysLys-----Pro 19
Db 6080 GAAATACTATAAAATAATACGACTCACTTTCGCGCTCGCAAAATACACTACTATAA 6021
QY 20 LysLysGlyAsnProAspGluGluArgLysArgGlnGlu-----MetThrThr 35
Db 6020 AAACAAATAACCCGATAAAAAAATAAATAAATAAATAAATAAATAAATAAATAA 5961
QY 36 PheGluArg-----LysLeuGlnAspGlnAspLysLysSer 47
Db 5960 TTATAAAATCTAACCCAAATCTAAAAACCCCTTAAATAACGAAAAAACTTAAAAATCT 5901
QY 48 GlnGluValSerSerThrSerAsnGlnGluAsnGluArgLysGlySerGlySerGluGluVal 67
Db 5900 AAAAACCCCTTACCAACCAATTAACCTTAAATAAATAAATAAATAAATAAATAA 5841
QY 68 CysTyrThr---ValIleAsnHisIleProHisGlnArgSerSerLeuSerSerAsnAsp 86
Db 5840 CTAATTCCTCTACTCACTATCTCGATACCCACGTAATAAATAAATAAATAAATAA 5781
QY 87 AspGlyTyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSer 106
Db 5780 CACAATATCTATAAAATTCATAATATTTTAAAAAAATCCGCTCTTACTAAAAACAAC 5721
QY 107 GluThrGluTyrAlaLeuLeuArgThrSerValSerArgProCysSerCysThrHisGlu 126
Db 5720 AAAACGACACCACTCGAAAAA-----AACCAATAATACACACGTTTC 5673
QY 127 HisAspTyrGluValValPhe 133
Db 5672 CACGATAAAAAAAATCTAT 5652

RESULT 12
US-10-261-175A-9
; Sequence 9, Application US/10261175A
; Publication No. US2004003822A1
; GENERAL INFORMATION:
; APPLICANT: DIETRICH, WILLIAM
; APPLICANT: WATERS, JAMES W.
; TITLE OF INVENTION: ANTHRAX SUSCEPTIBILITY GENE
; FILE REFERENCE: 56491(71250)
; CURRENT APPLICATION NUMBER: US/10/261,175A
; CURRENT FILING DATE: 2002-09-30
; PRIOR APPLICATION NUMBER: 60/325,864
; PRIOR FILING DATE: 2001-09-29
; NUMBER OF SEQ ID NOS: 23
; SOFTWARE: Patent Ver. 2.1
```



```
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 10372
; LENGTH: 1173
; TYPE: DNA
; ORGANISM: Human
US-10-027-632-10372

Alignment Scores:
Pred. No.: 0.0658      Length: 1173
Score: 90.00           Matches: 38
Percent Similarity: 40.27%      Conservative: 22
Best Local Similarity: 25.50%    Mismatches: 55
Query Match: 12.59%             Indels: 34
DB: 16                    Gaps: 6

US-10-005-907-2 (1-135) x US-10-027-632-10372 (1-1173)
QY 1 MetGlyLeuValCysThrValLeuArg-----LysLeuSerCysLeuGlyGlu 14
Db 542 ATGGCACTCCATCATGACGACCCCAATCTTCAGCTTCCAGTTTCGCCCATGCTTAC 601
QY 15 AsnGlnLysLysProLysLysGlyAsnProAsp-----GluGluArgLys 29
Db 602 CCTCAGATGAAGCCGAAGTACCTCACAAATCTCGTCTAGTTAACTTGACAGAGATCC 661
QY 30 ArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGlu 49
Db 662 CGTAYGAAGTACTGACATACCAATCTCTCTGCTGACCTGGAACTAACAGATCCTA 721
QY 50 ValSerSerThrSerAsnGlnGlu-----AsnGluAsnGlySerGlySer 64
Db 722 AATGACTCTAGTTCTGTATGAGAGAGTTAAATGAAATAGACAGTGAAGATGGT----- 772
QY 65 GluGluValCysThrThrValIleAsnHisIleProHisGlnArgSerLeuSerSer 84
Db 773 -----TTAAGTCTAGGATCACACAGATCAGGCCTGCTGCA 811
QY -----AsnAspAspGlyTyrGluAsnIleAspSerLeuThrArgLysValArg 100
Db 812 GAGCAGCTGATGGGCTCAGATGGCAACAATAATTGGAGACCAAGGGGATTCATTAGA 871
QY 101 GlnPheArgGluArgSerGluThrGluTyr-----AlaLeuLeuArgThrSer 116
Db 872 AGATTTCATGATAGTCTTCAGTGGCCCTTTTGCTCTTCTCACCATGATCGACGTAGC 931
QY 117 ValSerArgProCysSerCysThrHis 125
Db 932 ATCTCTCGTCACATAGAAAACATCCAC 958

RESULT 15
US-10-296-115-693
; Sequence 693, Application US/10296115
; Publication No. US20040053248A1
; GENERAL INFORMATION:
; APPLICANT: Hyseq Inc
; TITLE OF INVENTION: No. US20040053248A1el Nucleic Acids and Polypeptides
; FILE REFERENCE: 784PCT
; CURRENT APPLICATION NUMBER: US/10/296,115
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: US09/488,725
; PRIOR FILING DATE: 2000-01-21
; PRIOR APPLICATION NUMBER: US09/552,317
; PRIOR FILING DATE: 2000-04-25

; NUMBER OF SEQ ID NOS: 1478
; SEQ ID NO 693
; LENGTH: 1671
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-296-115-693

Alignment Scores:
Pred. No.: 0.109      Length: 1671
Score: 90.00           Matches: 37
Percent Similarity: 38.00%      Conservative: 20
Best Local Similarity: 24.67%    Mismatches: 62
Query Match: 12.59%             Indels: 31
DB: 13                    Gaps: 4

US-10-005-907-2 (1-135) x US-10-296-115-693 (1-1671)
QY 5 LeuLeuArgLysLeuSerCysLeuGlnGlnLysLysProLysLysGlyAsnPro 24
Db 889 CTCCTACAAATCCTCTTGTCCCTGGAGCGCAGGCAAGCGCCAGACCAAGCAGGAG 948
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 949 CAAGGAGTGGACACAGCAGGAGCGCCACACAGACACACAGCAGGAGAGAGGGCAGAAA 1008
QY 45 LysLysSerGlnGluValSerSerThrSerAsnGlnGlnAsnGluAsnGlySerGlySer 64
Db 1009 CAGGAAGAGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1068
QY 65 GluGlu-----ValCysTyrThrValIleAsnHisIleProHisGlnArgSer 80
Db 1069 AAGGAGGACGGAGAGGCTGTCTCAGCTGACAGACAGACTCAGAGCCCAAGTTTCACTCT 1128
QY 81 ----SerLeuSerSerAsnAspAspGlyTyr----- 89
Db 1129 GAATCTCTATCTTCTTAACCCCTTCTTGTCTCCCGGTACGAGAAAGTAGAGTCTACT 1188
QY 90 -----GluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGlu--- 104
Db 1189 CCTATGATATGAGAGAACATCCAGAGGCTCATTCATCAGCCAGGAATAGATGAAATG 1248
QY 105 -----ArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 1249 AATGAAATATATGATGAGAACTCTCTACTGGAGAAACCAAAACCCCTGGCAGGCTCTCGCAG 1308
QY 115 -ThrSerValSerArgProCysSerCys 123
Db 1309 CTGCCCCACACAGAGCCCTTGTCTGTGC 1336

Search completed: April 22, 2004, 19:43:59
Job time : 345 secs
```

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model
Run on: April 22, 2004, 17:08:28 ; Search time 81 Seconds
(without alignments)
924.918 Million cell updates/sec

Title: US-10-005-907-2
Perfect score: 715
Sequence: 1 MGNLLRLKSLCNGNQKPK.....SVSRPCSTHERDYEVVPPH 135

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 682709 seqs, 277475446 residues
Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n_model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US10005907/runat_21042004_113755_11461/app_query.fasta_1.327
-DB=issued_patents_NA -QFMI=fastap -SUFFIX=rni -MINMATCH=0.1 -LOPC=0
-LOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFT=pct -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USBR=US10005907 @CNG 1 1 85 @runat_21042004_113755_11461 -NCPUS=6 -ICPU=3
-MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/ptodata/2/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS_COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES						
Result No.	%		DB	ID	Description	
	Score	Query Match				
1	96	13.4	2301	1	US-08-306-691B-23	Sequence 23, Appl
2	96	13.4	2301	4	US-09-167-206-3	Sequence 3, Appl
3	96	13.4	2301	5	PCT-US93-06251-78	Sequence 78, Appl
4	90	12.6	8257	4	US-09-595-684B-30	Sequence 30, Appl
5	90	12.6	8503	4	US-09-620-312D-130	Sequence 130, App
6	89.5	12.5	1071	4	US-09-205-258-118	Sequence 118, App
7	86.5	12.1	1302	4	US-08-630-915A-17	Sequence 17, Appl
8	85.5	12.0	6755	3	US-08-931-999-4	Sequence 4, Appl
9	84	11.7	5394	3	US-08-688-376-1	Sequence 1, Appl
10	84	11.7	8146	4	US-09-976-594-725	Sequence 725, App
11	84	11.7	11917	4	US-09-566-921-32	Sequence 32, Appl
12	82.5	11.5	598	4	US-08-961-527-271	Sequence 271, App

13	82.5	11.5	6008	1	US-07-789-515A-5	Sequence 5, Appli
14	82.5	11.5	6008	1	US-08-005-002C-5	Sequence 5, Appli
15	82.5	11.5	6008	1	US-08-487-203A-5	Sequence 5, Appli
16	82	11.5	4262	4	US-09-521-511C-10	Sequence 10, Appl
17	81.5	11.4	563	4	US-09-621-976-2759	Sequence 2759, Ap
18	81	11.3	2885	4	US-10-164-595-39	Sequence 39, Appl
19	80	11.2	1428	2	US-08-786-399-2	Sequence 2, Appli
20	80	11.2	1428	2	US-08-989-387-8	Sequence 8, Appli
21	80	11.2	1428	4	US-09-216-387-2	Sequence 2, Appli
22	80	11.2	1494	1	US-08-585-758A-3	Sequence 3, Appli
23	80	11.2	1494	1	US-08-977-818-3	Sequence 3, Appli
24	80	11.2	1494	2	US-08-670-274B-3	Sequence 3, Appli
25	80	11.2	1494	3	US-09-146-187-3	Sequence 3, Appli
26	80	11.2	1494	3	US-09-033-655-997	Sequence 997, App
27	80	11.2	5761	4	US-09-620-312D-147	Sequence 147, App
28	80	11.2	5857	4	US-09-220-132-79	Sequence 79, Appl
29	80	11.2	6107	4	US-09-976-594-557	Sequence 557, App
30	79.5	11.1	512	4	US-09-621-976-3471	Sequence 3471, Ap
31	79.5	11.1	843	2	US-08-933-750C-67	Sequence 67, Appl
32	79.5	11.1	843	3	US-09-234-613-67	Sequence 67, Appl
33	79.5	11.1	4226	4	US-09-620-312D-480	Sequence 480, App
34	78.5	11.1	6002	4	US-09-345-882-4	Sequence 4, Appli
35	78.5	11.0	1491	3	US-08-999-774A-7	Sequence 7, Appli
36	78	10.9	1056	4	US-09-134-001C-1550	Sequence 1550, Ap
37	78	10.9	2139	3	US-09-026-343-1	Sequence 1, Appli
38	78	10.9	2139	4	US-09-362-871-1	Sequence 1, Appli
39	78	10.9	2791	4	US-09-570-367C-1	Sequence 1, Appli
40	78	10.9	2791	4	US-09-915-524-1	Sequence 1, Appli
41	77.5	10.8	2421	4	US-09-417-137-52	Sequence 52, Appl
42	77	10.8	2128	4	US-09-620-312D-197	Sequence 197, App
43	77	10.8	2623	4	US-09-540-824-23	Sequence 23, Appl
44	76.5	10.7	1238	4	US-09-183-861-75	Sequence 75, Appl
45	76.5	10.7	1238	4	US-09-022-765-75	Sequence 75, Appl

ALIGNMENTS

RESULT 1
US-08-306-691B-23
; Sequence 23, Application US/08306691B
; Patent No. 5734039
; GENERAL INFORMATION:
; APPLICANT: Calabretta, Bruno
; APPLICANT: Skorski, Tomasz
; TITLE OF INVENTION: ANTISENSE
; TITLE OF INVENTION: OLIGONUCLEOTIDES TARGETING COOPERATING ONCOGENES
; NUMBER OF SEQUENCES: 55
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Seidel, Gonda, Lavorigna & Monaco, P.C.
; STREET: Two Penn Center, Suite 1800
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: U.S.A.
; ZIP: 19102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette, 3.50 inch, 720 Kb
; COMPUTER: IBM PS/2
; OPERATING SYSTEM: MS-DOS
; SOFTWARE: WordPerfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/306.691B
; FILING DATE: September 15, 1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Monaco, Daniel A.
; REGISTRATION NUMBER: 30,480
; REFERENCE/DOCKET NUMBER: 8321-8
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (215) 568-8383
; TELEFAX: (215) 568-5549


```

; MOLECULE TYPE: DNA (genomic)
PCT-US93-06251-78

Alignment Scores:
Pred. No.: 0.0132 Length: 2301
Score: 96.00 Matches: 29
Percent Similarity: 41.41% Conservative: 24
Best Local Similarity: 22.66% Mismatches: 51
Query Match: 13.43% Indels: 24
DB: 5 Gaps: 3

US-10-005-907-2 (1-135) x PCT-US93-06251-78 (1-2301)
QY 14 GluAsnGlnLysLysProLysLysGlyAsnProAspGluGluArgLysArgGlnGluMet 33
Db 17 GAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 76
QY 34 ThrThrPheGluArgLysLeuGlnAspGlnAspLysLysSerGlnGluValSerSerThr 53
Db 77 GAGGAGCAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 136
QY 54 SerAsnGlnGluAsnGlnLysSerGlySerGluGluValCys----- 68
Db 137 GAGGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 196
QY 69 -----TyrThrValIleAsn 73
Db 197 AGGAACCGGAGCGGAGCAGCAGTAGCTGGTGGGCACCATGCTGGATCACCACCTCGAG 256
QY 74 HistleProHisGlnArgSerSerLeuSerSerAsn---AspAspGlyTyrGluAsnIle 92
Db 257 GCGTGAAGCGCAGATCCAGGTTCTCAGCAGCAGGAGCAGATGTCAGAGGAGCGAGCT 316
QY 93 AspSerLeuThrArgLysVal-----ArgGlnPheArgGluArgSerGluThrGlu 109
Db 317 GAGCGCTCCAGCGAGAAGTTGAGGAGAGAAGCGCGCGCGGAGCAGGCTGAGGCTGAG 376
QY 110 TyrAlaLeuLeuArgThrSerVal 117
Db 377 GTGGCTCTCTTGAACCGTAGGATC 400

RESULT 4
US-09-595-684B-30
; Sequence 30, Application US/09595684B
; Patent No. 6544766
; GENERAL INFORMATION:
; APPLICANT: Beraud, Christophe
; APPLICANT: Ohashi, Cara
; APPLICANT: Sakowicz, Roman
; APPLICANT: Vaisberg, Eugeni
; APPLICANT: Wood, Kenneth
; APPLICANT: Yu, Ming
; TITLE OF INVENTION: Human kinesins and methods of producing
; TITLE OF INVENTION: and purifying human kinesins
; FILE REFERENCE: cytop036
; CURRENT APPLICATION NUMBER: US/09/595,684B
; PRIOR FILING DATE: 2002-06-24
; PRIOR APPLICATION NUMBER: 09/295,612
; PRIOR FILING DATE: 2000-04-20
; NUMBER OF SEQ ID NOS: 105
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 30
; LENGTH: 8257
; TYPE: DNA
; ORGANISM: Human
US-09-595-684B-30

Alignment Scores:
Pred. No.: 0.551 Length: 8257
Score: 90.00 Matches: 25
Percent Similarity: 49.57% Conservative: 32
Best Local Similarity: 21.74% Mismatches: 44
Query Match: 12.59% Indels: 14

DB: 4 Gaps: 4
US-10-005-907-2 (1-135) x US-09-595-684B-30 (1-8257)
QY 5 LeuLeuArgLysLeuSerCysLeuGluAsnGlnLysLysProLysLysGlyAsnPro 24
Db 3433 CTTTCTAGGACCTGTGACAGACTGGCAGAGTTGAGAAAGAACTAAAGGAAAGAGCCAG 3492
QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
Db 3493 CAACTCCCAAGAAACACAGCAACTTCTTAAATGACAAAGAGAGATGAGTGCAGTGCAG 3552
QY 45 LysLysSerGlnGluValSerSerThrSerAsnGln---GluAsnGluAsnGlySerGly 63
Db 3553 AAAAAGATTATGAATAGAGAAATTTAAAGATGAATTAAGAAACAAGAA----- 3603
QY 64 SerGluGluValCysTyrThrValIleAsnHistleProHisGlnArgSerSerLeuSer 83
Db 3604 -----TTGACATTGGAAACATATGGAACACAGAGAGCGCTTGAGTTGGCT 3645
QY 84 SerAsn---AspAspGlyTyrGluAsnIleAspSerLeuThrArgLys-----Val 99
Db 3646 CAGAACTTAATGAATTTATGAGGAAGTGAATCTATTAACCAAGAAAGAAAGTTCTTA 3705
QY 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArg 114
Db 3706 AAGGAATTACAGAGTCACTTTTGAACACAGAGAGACCCCTTAGA 3750

RESULT 5
US-09-620-312D-130
; Sequence 130, Application US/09620312D
; Patent No. 6569662
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Zhang, Jie
; APPLICANT: Ren, Feiyan
; APPLICANT: Chen, Rui-hong
; APPLICANT: Zhao, Qing A.
; APPLICANT: Wehrman, Tom
; APPLICANT: Xue, Aidong J.
; APPLICANT: Yang, Yonghong
; APPLICANT: Wang, Jian-Rui
; APPLICANT: Zhou, Ping
; APPLICANT: Ma, Yuning
; APPLICANT: Wang, Dunrui
; APPLICANT: Wang, Zhiwei
; APPLICANT: John Tillinghast
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. 6569662el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 784CIP28
; CURRENT APPLICATION NUMBER: US/09/620,312D
; CURRENT FILING DATE: 2000-07-19
; PRIOR APPLICATION NUMBER: 09/552,317
; PRIOR FILING DATE: 2000-04-25
; PRIOR APPLICATION NUMBER: 09/488,725
; PRIOR FILING DATE: 2000-01-21
; NUMBER OF SEQ ID NOS: 1105
; SOFTWARE: pt_FL_genes Version 1.0
; SEQ ID NO 130
; LENGTH: 8503
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (91)..(8082)
US-09-620-312D-130

Alignment Scores:
Pred. No.: 0.576 Length: 8503
Score: 90.00 Matches: 25

```

Percent Similarity: 49.57% Conservative: 32
 Best Local Similarity: 21.74% Mismatches: 44
 Query Match: 12.53% Indels: 14
 DB: 4 Gaps: 4

US-10-005-907-2 (1-135) x US-09-620-312D-130 (1-8503)

QY 5 LeuLeuArgLysLeuSerCysLeuGluGlyGlnLysProLysLysGlyAsnPro 24
 DB 3433 CTTTCTAGGACCTGGACAGACTGGCAGAGTTGAAGAAAACCTAAAGGAAGAGCCAG 3492
 QY 25 AspGluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAsp 44
 DB 3493 CAATCCCAAGAAAACACCAACAACCTTTCTTAATGTACAAGAGAGATGATGAGATCCAG 3552
 QY 45 LysLysSerGlnGluValSerThrSerAsnGln---GluAsnGluAsnGlySerGly 63
 DB 3553 AAAAGATTATGAATAGAGAAATTAAGATGAATTAAGACAAGAA----- 3603
 QY 64 SerGluGluValCysTyrThrValIleAsnHisIleProHisGlnArgSerSerLeuSer 83
 DB 3604 -----TTGACATTGGACATATATGGAACACAGAGAGCGCTTGAGTTGGCT 3645
 QY 84 SerAsn---AspAspGlyTyrGluAsnIleAspSerLeuThrArgLys-----Val 99
 DB 3646 CAGAAACTTAATGAAATATGAGAGAGTGAATCTATACCAAGAAAGAAAAGTTCTA 3705
 QY 100 ArgGlnPheArgGluArgSerGluThrGluTyrAlaLeuLeuArg 114
 DB 3706 AAGGAATTACAGAAAGTCAATTTGAAACACAGAGAGACCACCTTAGA 3750

RESULT 6
 US-09-205-258-118
 ; Sequence 118, Application US/09205258
 ; Patent No. 6525174
 ; GENERAL INFORMATION:
 ; APPLICANT: Young et al.
 ; TITLE OF INVENTION: 207 Human Secreted Proteins
 ; FILE REFERENCE: PZ007P1
 ; CURRENT APPLICATION NUMBER: US/09/205,258
 ; CURRENT FILING DATE: 1998-12-04
 ; EARLIER APPLICATION NUMBER: PCT/US98/11422
 ; EARLIER FILING DATE: 1998-06-04
 ; EARLIER APPLICATION NUMBER: 60/048,885
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,375
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,881
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,880
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,896
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,020
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,876
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,895
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,884
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,894
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,971
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,964
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,882
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,899
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,893
 ; EARLIER FILING DATE: 1997-06-06

; EARLIER APPLICATION NUMBER: 60/048,900
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,901
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,892
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,915
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,019
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,970
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,972
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,916
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,373
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,875
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/049,374
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,917
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,949
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,974
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,883
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,897
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,898
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,962
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,963
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,877
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/048,878
 ; EARLIER FILING DATE: 1997-06-06
 ; EARLIER APPLICATION NUMBER: 60/070,923
 ; EARLIER FILING DATE: 1997-12-18
 ; EARLIER APPLICATION NUMBER: 60/092,921
 ; EARLIER FILING DATE: 1998-07-15
 ; EARLIER APPLICATION NUMBER: 60/094,657
 ; EARLIER FILING DATE: 1998-07-30
 ; NUMBER OF SEQ ID NOS: 1227
 ; SOFTWARE: Patent in Ver. 2.0
 ; SEQ ID NO 118
 ; LENGTH: 1071
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-205-258-118

Alignment Scores:
 Pred. No.: 0.0303 Length: 1071
 Score: 89.50 Matches: 41
 Percent Similarity: 36.26% Conservative: 21
 Best Local Similarity: 23.98% Mismatches: 56
 Query Match: 12.52% Indels: 53
 DB: 4 Gaps: 7

US-10-005-907-2 (1-135) x US-09-205-258-118 (1-1071)

QY 10 SerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnProAspGluArgLys 29
 DB 47 TCCATCTCTGAGCTCCATGGAGAGCCACCAGCTCGGTGACCAGGAGACTCGGCGCAAG 106
 QY 30 ArgGlnGluMetThrThrPheGluArgLysLeuGlnAspGlnAspLysSerGlnGlu 49
 DB 107 GCCCGAGACAGCGCCGCCCTGGAAGAACTACAAGACAGAGAAACAA---CAGAAA 163

ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6673549 4295277CB1
NAME/KEY: unsure
LOCATION: 8083
OTHER INFORMATION: a, t, c, g, or other
US-09-976-594-725

Alignment Scores:
Pred. No.: 3.35 Length: 8146
Score: 84.00 Matches: 32
Percent Similarity: 38.02% Conservatives: 14
Best Local Similarity: 26.45% Mismatches: 47
Query Match: 11.75% Indels: 28
DB: 4 Gaps: 5

US-10-005-907-2 (1-135) x US-09-976-594-725 (1-8146)

QY 13 GlyGluAsnGlnLysLysPro-----LysLysGlyAsnProAsp 25
DB 2886 GCGGATGTGAACAGACACCTTGAGGAGCTGGAGCCCGTCGAGAGGAGGAGGTAGACGAC 2945
QY 26 GluGluArgLysArgGlnGluMetThrPheGluArgLysLeuGlnAspGlnAspLys 45
DB 2946 ATTGAATAATTGAAGATGAAGGAGCGCGTTTTCAGAGATCTTCAGAGACTGGAGAC--- 3002
QY 46 LysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySerGlySerGlu 65
DB 3003 ---TATGAAGAGAGGCGGAACTGAGGAGGCTGAGGAGCCAGAGAGGATGGGAGGAA 3059
QY 66 GluValCysTyrThrValIleAsnHisIlePro----- 76
DB 3060 CACGTATGTGTGAGCGGCTCCAGACACAGCCCACTGAGGATGAGGAAAGTGCAGAGCG 3119
QY 77 -----HisGlnArgSerSerLeuSerSerAsnAspAsp---Gly 88
DB 3120 GAGCTGATGATCATCATCAGGAGAGGAGGAGTCTGTGGCCATGGGATGACCGGCC 3179
QY 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108
DB 3180 GAAGAGACATGATGATGAGCCATTCAGGAGGAGGAG-----GCTGAACAATCTGAAGAG 3233
QY 109 Glu 109
DB 3234 GAG 3236

RESULT 11
US-09-566-921-32
Sequence 32, Application US/09566921
Patent No. 6882888
GENERAL INFORMATION:
APPLICANT: Loring, Jeanne F.
APPLICANT: Tingley, Debora W.
APPLICANT: Edwards, Carla M.
TITLE OF INVENTION: GENES EXPRESSED IN ALZHEIMER'S DISEASE
FILE REFERENCE: PA-0024 US
CURRENT APPLICATION NUMBER: US/09/566,921
CURRENT FILING DATE: 2000-05-05
NUMBER OF SEQ ID NOS: 138
SOFTWARE: PERL Program
SEQ ID NO 32
LENGTH: 11917
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
OTHER INFORMATION: Incyte ID No. 6682888 475473.1
NAME/KEY: unsure
LOCATION: 9384, 10290-10321, 10325
OTHER INFORMATION: a, t, c, g, or other
US-09-566-921-32

Alignment Scores:
Pred. No.: 5.92 Length: 11917
Score: 84.00 Matches: 32
Percent Similarity: 38.02% Conservatives: 14
Best Local Similarity: 26.45% Mismatches: 47
Query Match: 11.75% Indels: 28
DB: 4 Gaps: 5

US-10-005-907-2 (1-135) x US-09-566-921-32 (1-11917)

QY 13 GlyGluAsnGlnLysLysPro-----LysLysGlyAsnProAsp 25
DB 2887 GCGGATGTGAACAGACACCTTGAGGAGCTGGAGCCCGTCGAGAGGAGGAGGTAGACGAC 2946
QY 26 GluGluArgLysArgGlnGluMetThrPheGluArgLysLeuGlnAspGlnAspLys 45
DB 2947 ATTGAATAATTGAAGATGAAGGAGCGCGTTTTCAGAGATCTTCAGAGACTGGAGAC--- 3003
QY 46 LysSerGlnGluValSerSerThrSerAsnGlnGluAsnGlySerGlySerGlu 65
DB 3004 ---TATGAAGAGAGGCGGAACTGAGGAGGCTGAGGAGCCAGAGAGGATGGGAGGAA 3060
QY 66 GluValCysTyrThrValIleAsnHisIlePro----- 76
DB 3061 CACGTATGTGTGAGCGGCTCCAGACACAGCCCACTGAGGATGAGGAAAGTGCAGAGCG 3120
QY 77 -----HisGlnArgSerSerLeuSerSerAsnAspAsp---Gly 88
DB 3121 GAGCTGATGATCATCATCAGGAGAGGAGGAGTCTGTGGCCATGGGATGACCGGCC 3180
QY 89 TyrGluAsnIleAspSerLeuThrArgLysValArgGlnPheArgGluArgSerGluThr 108
DB 3181 GAAGAGACATGATGATGAGCCATTCAGGAGGAGGAG-----GCTGAACAATCTGAAGAG 3234
QY 109 Glu 109
DB 3235 GAG 3237

RESULT 12
US-08-961-527-271/c
Sequence 271, Application US/08961527
Patent No. 6420135
GENERAL INFORMATION:
APPLICANT: Charles Kunsch
TITLE OF INVENTION: Streptococcus pneumoniae Polynucleotides and Sequences
NUMBER OF SEQUENCES: 391
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.4Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MSDOS version 6.2
SOFTWARE: ASCII Text
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/961,527
FILING DATE:
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Brookes, A. Anders
REGISTRATION NUMBER: 36,373
REFERENCE/DOCKET NUMBER: PB340P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (301) 309-8504
TELEFAX: (301) 309-8512
INFORMATION FOR SEQ ID NO: 271:

SEQUENCE CHARACTERISTICS:
 LENGTH: 598 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: linear
 US-08-961-527-271

Alignment Scores:

Pred. No.: 0.107
 Score: 82.50
 Percent Similarity: 41.71%
 Best Local Similarity: 27.27%
 Indels: 11.54%
 Gaps: 4

US-10-005-907-2 (1-135) x US-08-961-527-271 (1-598)

```

QY 3 AsnTyr-----LeuLeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLys 18
Db 534 AACTACAGATTAGTACAGCTTACGAGAGAAAGCTGATTGCCAGACACCTCAAGAA 475
QY 19 ProLysLysGlyAsnProAspGluArgLysArgGlnGluMetThrThrPheGluArg 38
Db 474 CCGAARAAG-----TCCAGAAG-----TCATCGAGTTCGCAG 442
QY 39 LysLeuGlnAspGlnAspLysLysSerGlnGluValSer----- 51
Db 441 AAATCTCGACCAAGAGAAATAATCAAGAGAAATCGATGATCATCGTCTATCTAGTTG 382
QY 52 -----SerThrSer-AsnGlnGluAsnGlnGly----- 61
Db 381 AGATTGGCAAGATTACGGCTTCTATATCGAGATCGAGAGTTCAGCAGAAACCGCA 322
QY 62 -----SerGlySerGluGluValCysTyrThrVa 71
Db 321 GTCAGACCTACGACGATTGATTCGCACTCTATCAGGTTCTTA-AACTCTTTGCCAACCAAT 263
QY 71 lileasnHisleProHisGlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAs 91
Db 262 TTTATCCAAATTCCTCAATCAGAAATCATCATATCG----- 226
QY 91 nileaspSerLeuThrArgLysValArgGlnPhe-----Ar 103
Db 225 -ATTCCATCTCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 167
QY 103 gLuarGSerGluThrGluTyrAlaLeuLeu----- 113
Db 166 GAAAGCTTCACTACTCTTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 107
QY 114 ----ArgThrSerValSerArgProCysSer-----CysThrHisGlu---HisAspTy 129
Db 106 TGAAAAAACAAAGTATTTCACGA---TGCAGCTACCGGTGTTCTTATCGGCTCATCCTCT 50
QY 129 rGluValValPheProHis 135
Db 49 CTATCCTCTTTTCACTCAT 31

```

RESULT 13

US-07-789-915A-5

Sequence 5, Application US/07789915A

Patent No. 5212059

GENERAL INFORMATION:

APPLICANT: Baker, Rohan T.

APPLICANT: Tobias, John W.

APPLICANT: Varshavsky, Alexander

TITLE OF INVENTION: Ubiquitin-Specific Proteases

NUMBER OF SEQUENCES: 8

CORRESPONDENCE ADDRESSES:

ADDRESS: Hamilton, Brook, Smith & Reynolds, P.C.

STREET: Two Militia Drive

CITY: Lexington

STATE: Massachusetts

COUNTRY: U.S.A.

ZIP: 02173
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.25
 CURRENT APPLICATION DATA: US/07789,915A
 APPLICATION NUMBER: 07/789,915A
 FILING DATE: 19911108
 CLASSIFICATION: 435

ATTORNEY/AGENT INFORMATION:

REGISTRATION NUMBER: 22,592

REFERENCE/DOCKET NUMBER: MIT-5091AA

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617-861-6240

TELEFAX: 617-861-9540

NAME: Brook, David E.

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 6008 base pairs

TYPE: NUCLEIC ACID

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

NAME/KEY: CDS

LOCATION: 983..4774

US-07-789-915A-5

Alignment Scores:

Pred. No.: 3.36
 Score: 82.50
 Percent Similarity: 36.80%
 Best Local Similarity: 25.60%
 Query Match: 11.54%
 Indels: 20
 Gaps: 3

US-10-005-907-2 (1-135) x US-07-789-915A-5 (1-6008)

```

QY 6 LeuArgLysLeuSerCysLeuGlyGluAsnGlnLysLysProLysLysGlyAsnProAsp 25
Db 5112 ATGAGAAACAAAGCCGCTCTCAA-GATGATAATGTCAAGGAAGTCAATGGAATGATGAT 5170
QY 26 GluGluArgLysArgGlnGluMetThrThrPheGluArgLysLeu----- 40
Db 5171 ACTAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5230
QY 41 -----GlnAspGlnAspLysLysSerGlnGluValSerSerThrSerAsnGlnGlu 57
Db 5231 GATGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5290
QY 58 AsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHis 77
Db 5291 GATGAGAGACGACGAGGAGAGAAAT-----ATCACCCATAGAAAGTCGT 5332
QY 78 GlnArgSerSerLeuSerSerAsnAspAspGlyTyrGluAsnIleAspSerLeuThrArg 97
Db 5333 CATAGAGAAAGTCTCGGTTAGATGATGATGATGATGATGATGATGATGATGATGATGAT 5392
QY 98 Lys-----ValArgGlnPheArgGluArgSerGluThrCluTyrAlaLeu 112
Db 5393 AAAATTTTATGTTGAGGATGCTAATAACTTTTCTGATCTTGTGATGAGACTACTCACACT 5452
QY 113 LeuArgThrSerVal 117
Db 5453 AAAGAGAAATGTA 5467

```

RESULT 14

US-08-005-002C-5

Sequence 5, Application US/08005002C

Patent No. 5494818

GENERAL INFORMATION:

APPLICANT: Baker, Rohan T.

APPLICANT: Tobias, John W.


```
Qy 58 AsnGluAsnGlySerGlySerGluGluValCysTyrThrValIleAsnHisIleProHis 77
Db 5291 GATGAGACGACGAGGAGAGAAAT-----ATCAACCATAGAGATCGT 5332
Qy 78 GlnArgSerSerLeuSerSerAsnAspGlyTyrGluAsnIleAspSerLeuThrArg 97
Db 5333 CATAGAGAACTCTCGGGTTAGATGATGACGACGAGGAGCAAGCCATGTACACCCGA 5392
Qy 98 Lys-----ValArgGlnPheArgGluArgSerGluThrGluTyrAlaLeu 112
Db 5393 AAATTTTATGGTGAGGATGCTATAACTTTTCTGATCTTGATGAGACTACTCACACTTTT 5452
Qy 113 LeuArgThrSerVal 117
Db 5453 AAAGAGGAAATGTA 5467
```

Search completed: April 22, 2004, 18:40:34
Job time : 99 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 11:22:40 ; Search time 351 Seconds
(without alignments)
4901.767 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405
Sequence: 1 atgggaattatctctctg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N: Geneseq_29Jan04:*

1: Geneseqn1980s:*

2: Geneseqn1990s:*

3: Geneseqn2000s:*

4: Geneseqn2001as:*

5: Geneseqn2001bs:*

6: Geneseqn2002as:*

7: Geneseqn2003as:*

8: Geneseqn2003bs:*

9: Geneseqn2003cs:*

10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	100.0	667	9	ADC30760 Human nov
2	405	100.0	3762	6	ABN81319 Human mas
3	58	14.3	732	7	ACC59951 Human IRA
4	47.2	11.7	1643	6	ABL99791 Human sec
5	46.6	11.5	775	7	ABX05108 Human nov
6	41.6	10.3	6446	4	AS46328 Tumour su
7	41.6	10.3	18988	4	AA54341 Tumour su
8	41.6	10.3	18988	6	ABL32700 Human imm
9	41.6	10.3	18988	6	ABL34508 Human met
10	41.6	10.3	18988	6	ABL70203 Chemocall
11	41	10.1	1145	4	AAK17837 Human bra
12	41	10.1	1996	4	AAK05119 Human bra
13	41	10.1	11924	9	ADC30201 Human nov
14	41	10.1	11951	9	ADC30200 Human nov
15	41	10.1	14994	7	ADL19440 CDNA enco
16	40.4	10.0	6523	6	ABN80121 Human che
17	40.4	10.0	6523	9	ABE84216 Human lym
18	40.4	10.0	6523	9	ABE84140 Human lym
19	40.2	9.9	4186	2	AZ41320 Human nor
20	40	9.9	2921	3	AAC77829 Human can
21	40	9.9	3402	4	AAH54452 S. epider
22	40	9.9	3668	4	AAH54402 S. epider
23	40	9.9	4100	4	AAH54798 S. epider

ALIGNMENTS

RESULT 1

ADC30760

ID ADC30760 standard; CDNA; 667 BP.

XX

AC ADC30760;

XX

DT 18-DEC-2003 (first entry)

XX

DE Human novel cDNA sequence, SEQ ID NO:842.

XX

KW Human; diagnostic; drug screening; forensics; gene mapping;

KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;

KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;

KW ulcers; osteoporosis; autoimmune disease; cancer;

KW molecular weight marker; food supplement; antiparkinsonian; nootropic;

KW neuroprotective; antihaemic; anticoagulant; thrombolytic; vulnary;

KW antiulcer; osteopathic; immunosuppressive; antiinflammatory; cytostatic;

KW gene therapy; chromosome 1; gene; ss.

XX

OS Homo sapiens.

XX

PN WO2003029271-A2.

XX

PD 10-APR-2003.

PF 24-SEP-2002; 2002WO-US030474.

XX

PR 24-SEP-2001; 2001US-0324631P.

XX

PA (HYSE-) HYSEQ INC.

XX

PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;

PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;

XX

XX Haley-Vicente D, Drmanac RT;

XX WPI; 2003-371981/35.

DR P-PSDB; ADC31731.

XX

PT New polynucleotide and polypeptide useful for diagnosing, preventing or

PT treating conditions such as neurodegenerative diseases, anemias, platelet

PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or

PT cancer.

XX

PS Claim 1; SEQ ID NO 842; 1185pp; English.

XX

CC The invention relates to 971 novel human cDNA sequences (ADC29919-

CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 667 BP; 216 A; 147 C; 154 G; 150 T; 0 U; 0 Other;

Query Match 100.0%; Score 405; DB 9; Length 667;
 Best Local Similarity 100.0%; Pred. No. 2.9e-107;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 60
 DB 186 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 245
 QY 61 AAAGGAACCCAGATGAGGAAGAAACCGGAGGAATGACTACATTTGAAAGAACTT 120
 DB 246 AAAGGAACCCAGATGAGGAAGAAACCGGAGGAATGACTACATTTGAAAGAACTT 305
 QY 121 CAAGATCAGATACAAAGCCAGAGAGTTTCATCCACTTCTATCATCAGGAACGAGAT 180
 DB 306 CAAGATCAGATACAAAGCCAGAGAGTTTCATCCACTTCTATCATCAGGAACGAGAT 365
 QY 181 GGCAGTGGTTCTGAGAGAGTGTGCTACACTGTCTATTAATCAATCCATCCATCAGATCC 240
 DB 366 GGCAGTGGTTCTGAGAGAGTGTGCTACACTGTCTATTAATCAATCCATCCATCAGATCC 425
 QY 241 TCCCTGAGCTCCATGATGATGGCTATGAGACATTCCTCCCTCACAAGGAAGTGA 300
 DB 426 TCCCTGAGCTCCATGATGATGGCTATGAGACATTCCTCCCTCACAAGGAAGTGA 485
 QY 301 CAGTTTAGAAGAGTCCAGACAGAAATATGCCCTTCTTAGGACTTCTGTTAGTGGCCT 360
 DB 486 CAGTTTAGAAGAGTCCAGACAGAAATATGCCCTTCTTAGGACTTCTGTTAGTGGCCT 545
 QY 361 TGTTCCTGACCCATGACATGATATGAAGTTGTTTCACAC 405
 DB 546 TGTTCCTGACCCATGACATGATATGAAGTTGTTTCACAC 590

RESULT 2
 ABN81319
 ID ABN81319 standard; cDNA; 3762 BP.

XX
 AC
 XX

DT 30-AUG-2002 (first entry)
 XX Human mast cell related gene MC1 SEQ ID NO 1.
 DE
 XX Human; mast cell; MC; antiallergic; antiinflammatory; antiasthmatic;
 KW vasotropic; dermatological; allergy; hypersensitivity; rhinitis; asthma;
 KW gene; ss.
 XX Homo sapiens.
 OS
 XX Key Location/Qualifiers
 FT 25..432
 CDS /*tag= a
 /product= "MC1"
 XX WO200246389-A2.
 PN 13-JUN-2002.
 PD 07-DEC-2001; 2001WO-US046180.
 XX 08-DEC-2000; 2000US-0251835P.
 PR 14-MAR-2001; 2001US-0275479P.
 PR 28-MAR-2001; 2001US-0279115P.
 PR 02-APR-2001; 2001US-0280143P.
 XX (UNIO) UCB SA.
 PA Nocka K, Pirozzi G, Einstein R;
 FI WPI; 2002-508560/54.
 XX P-PSDB; ABB77569.
 DR
 XX Novel isolated nucleic acids that are differentially expressed in mast
 PT cells in patients with allergic hypersensitivity, encoding proteins
 PT associated with mast cell regranulation and allergic hypersensitivity.
 XX
 PS Claim 1; Page 95-97; 119pp; English.
 XX The invention relates to isolated nucleic acid (ABN81319-ABN81324),
 CC corresponding to genes differentially expressed in mast cells following
 CC activation or in patients with allergic hypersensitivity disease. (I)
 CC that encodes proteins (ABB77569-ABB77575) (II) or a protein fragment of
 CC (II) if at least 6 amino acids. (III) is useful for identifying binding
 CC partners. (I) or (II) is useful for diagnosing or treating a disease
 CC state (e.g. allergic hypersensitivity, seasonal rhinitis, asthma,
 CC urticaria or atopic dermatitis or mastocytosis) in a subject which
 CC involves determining the level of expression of (I) or (II). A computer
 CC system, comprising a database containing information identifying the
 CC expression level in a tissue or at least one mast cell of (I), is useful
 CC for presenting information to identify the relative expression level of
 CC (I). (II) is used as a marker to detect, diagnose or identify an allergic
 CC response in a patient. The protein can also serve as a target that
 CC modulate gene expression or activity and as an antigen to raise
 CC polyclonal or monoclonal antibodies. (II) is useful for identifying
 CC agents that modulate expression of the protein or agents, such as
 CC agonists or antagonists. The agonists or antagonists are useful for
 CC modulating biological activity and function of (II) and thus are useful
 CC for alleviating disease conditions such as allergic hypersensitivity,
 CC seasonal rhinitis, asthma, urticaria, atopic dermatitis or mastocytosis
 XX
 SQ Sequence 3762 BP; 1220 A; 672 C; 680 G; 1190 T; 0 U; 0 Other;
 Query Match 100.0%; Score 405; DB 6; Length 3762;
 Best Local Similarity 100.0%; Pred. No. 5.4e-107;
 Matches 405; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 60
 DB 25 ATGGGAATTAATCTCTCGGAAACTCAGTTGCTCGGAGAGATCAAAAGAGCCCAAG 84
 QY 61 AAAGGAACCCAGATGAGGAAGAAACCGGAGGAATGACTACATTTGAAAGAACTT 120

AC ABL99791;

03-OCT-2002 (first entry)

Human secretory polynucleotide (sptm) 46.

Human; ss; gene; secretory protein; secretory polynucleotides; SPM; SPTM-related disease; somatic gene therapy; germline gene therapy; severe combined immunodeficiency; intracellular parasite protection; fungal parasite; protozoan parasite; cell proliferative disorder; cancer; immune disorder; AIDS; neurological disorder; Parkinson's disease; motor neuron disorder; demyelinating disease; multiple sclerosis; meningitis; abscess; prion diseases; cerebral palsy; neuroskeletal disorder; peripheral nervous system disorder; dermatomyositis; polymyositis; myopathy; myasthenia gravis; mental disorder; Tourette's syndrome.

Homo sapiens.

XX PN WO200220756-A2.

XX
PD
14-MAR-2002.

XX
PF 30-AUG-2001: 2001WO-US027297.

05-SEP-2000: 2000US-0229747P.

05-SEP-2000; 2000US=0229749P.
PR

PR 05-SEP-2000; 2000US-0229750P.
PR 05-SEP-2000; 2000US-0229751P.

05-SEP-2000; 2000US=0230016F.
05-SEP-2000; 2000US=0230583P.
05-SEP-2000; 2000US=0230583P.
05-SEP-2000; 2000US=0230583P.

06-SEP-2000; 2000US=0230505A.
PR
06-SEP-2000; 2000US=0230514B.
PR

PR 06-SEP-2000; 2000US-0230515P.
PR 06-SEP-2000; 2000US-0230517P.

PR 06-SEP-2000; 2000US-0230518P.
PR 06-SEP-2000; 2000US-0230519P.

06-SEP-2000; 2000US-0230595P.
 PP
 06-SEP-2000; 2000US-0230596P.
 PP

06-SEP-2000; 2000US-0230597P.
06-SEP-2000; 2000US-0230598P.
06-SEP-2000; 2000US-0230599P.

PR 06-SEP-2000; 2000US-0230610P.
PR 06-SEP-2000; 2000US-0230864P.

PR 06-SEP-2000; 2000US-0230865P.

06-SEP-2000; 2000US-0230989P.

07-SEP-2000; 2000US-0230896P.

PR 07-SEP-2000; 2000US-0230951P.

PR 07-SEP-2000; 2000US-0231832P.

PA (INCY-) INCYTE GENOMICS INC.

PI Stuart J, Lincoln SE, Altus

PI Momiyama MG, Bradley DL, Ro

PI Marwaha R, Chen AJ, Chang S

DR WPI; 2002-315658/35.

XX

PT gene therapy of e.g. genetic

[illegible]

100

CC secretory (SPTM) proteins. The

10

AC	ABL99791;	
XX		
XX	03-OCT-2002 (first entry)	
DT		
XX		
DE	Human secretory polynucleotide (sptm) 46.	
XX		
KW	Human; ss; gene; secretory protein; secretory polynucleotides; SPTM;	
KW	SPTM-related disease; somatic gene therapy; germline gene therapy;	
KW	severe combined immunodeficiency; intracellular parasite protection;	
KW	fungus parasite; protozoan parasite; cell proliferative disorder; cancer;	
KW	immune disorder; AIDS; neurological disorder; Parkinson's disease;	
KW	motor neuron disorder; demyelinating disease; multiple sclerosis;	
KW	meningitis; abscess; prion diseases; cerebral palsy;	
KW	neurological disorder; peripheral nervous system disorder;	
KW	dermatomyositis; polymyositis; myopathy; myasthenia gravis;	
KW	mental disorder; Tourette's syndrome.	
XX		
OS	Homo sapiens.	
XX		
XX	WO200220756-A2.	
PN		
XX		
PD	14-MAR-2002.	
XX		
XX	30-AUG-2001; 2001WO-US027297.	
PF		
XX		
PR	05-SEP-2000; 2000US-0229747P.	
PR	05-SEP-2000; 2000US-0229748P.	
PR	05-SEP-2000; 2000US-0229749P.	
PR	05-SEP-2000; 2000US-0229750P.	
PR	05-SEP-2000; 2000US-0229751P.	
PR	05-SEP-2000; 2000US-0230016P.	
PR	05-SEP-2000; 2000US-0230583P.	
PR	05-SEP-2000; 2000US-0230505P.	
PR	06-SEP-2000; 2000US-0230514P.	
PR	06-SEP-2000; 2000US-0230515P.	
PR	06-SEP-2000; 2000US-0230517P.	
PR	06-SEP-2000; 2000US-0230518P.	
PR	06-SEP-2000; 2000US-0230519P.	
PR	06-SEP-2000; 2000US-0230596P.	
PR	06-SEP-2000; 2000US-0230597P.	
PR	06-SEP-2000; 2000US-0230599P.	
PR	06-SEP-2000; 2000US-0230610P.	
PR	06-SEP-2000; 2000US-0230864P.	
PR	06-SEP-2000; 2000US-0230865P.	
PR	06-SEP-2000; 2000US-0230988P.	
PR	06-SEP-2000; 2000US-0230989P.	
PR	07-SEP-2000; 2000US-0230896P.	
PR	07-SEP-2000; 2000US-0230897P.	
PR	07-SEP-2000; 2000US-0230951P.	
PR	07-SEP-2000; 2000US-0231163P.	
PR	07-SEP-2000; 2000US-02311832P.	
XX		
XX	(INCY-) INCYTE GENOMICS INC.	
XX		
XX	Stuart J, Lincoln SE, Altus CM, Dufour GE, Chalup MS, Hillman JL;	
PI	Jones AL, Yu JY, Wright RJ, Gietzen D, Liu TF, Yap PE, Dahl CR;	
PI	Moniyama MG, Bradley DL, Rohatgi SD, Harris B, Roseberry AM;	
PI	Gersting EH, Peralica CH, David MH, Panzer SR, Flores V, Daffo A;	
PI	Marwaha R, Chen AU, Chang SC, Au AP, Inman RR;	
XX		
XX	WPI; 2002-315658/35.	
DR	P-PSDB; ABB97794.	
DR		
XX		
XX	Polynucleotide sequences encoding human secretory proteins useful for	
XX	gene therapy of e.g. genetic deficiency disorders, cancers, and diseases	
XX	caused by intracellular parasites.	
XX		
XX	Claim 1; Page 275-276; 585pp; English.	
XX		
XX	The invention comprises the amino acid and coding sequences of human	
XX	secretory (SPTM) proteins. The SPTM DNA and amino acid sequences are	
CC		

useful for treating a disease or condition associated with the expression of functional SPTW. The SPTW DNA sequences are useful for somatic or germline gene therapy to correct a genetic deficiency (e.g. severe combined immunodeficiency). The SPTW DNA sequences are also useful in providing protection against intracellular parasites (e.g. fungal parasites and protozoan parasites). The SPTW DNA and protein sequences are also useful for diagnosing cell proliferative disorders, cancer, immune disorders (e.g. AIDS), neurological disorders (e.g. Parkinson's disease), motor neuron disorders, demyelinating diseases (e.g. multiple sclerosis), meningitis, abscesses, prion diseases, cerebral palsy, neuroskeletal disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, myopathy, myasthenia gravis, and mental disorders (e.g. Tourette's syndrome). cDNA sequences AB199746 - ABL99929 represent human secretory polynucleotides of the invention

Sequence 1643 BP; 413 A; 382 C; 384 G; 460 T; 0 U; 4 Other;

Query Match 11.7%; Score 47.2; DB 6; Length 1643;
Best Local Similarity 50.8%; Pred. No. 0.0024;
Matches 165; Conservative 0; Mismatches 154; Indels 6; Gaps 2;

87 ACGGAGGAAATGACTACATTTTGAAAGAAAATTTCAAGATCAAGTATAGAAAGCCAGA 146
1236 ATGGAAAAAATACTCATTTTGAAGAGGCAAGATTTCCCAAAACGAAATGAAGAAT 1295
147 AGTTTCATCCACTTCTTAATCAGGAAAACAGAAATGCGAGTGGTTCTGAAAGAGTGTGCTA 206
1296 GTCATCTACTCTCCATCCAGCAGGACCAATGTTGCCAGACCTACTCAGAGGAGCTGTGCTA 1355
207 CACTGTCAATTAATCA---CATCCCATCAGAGATCTCCCTGAGCTCCAATGATGATGG 263
1356 TACCCTCATCAATCATCGGGTCTCTGTTACAAGGCCATCAGGGAATCTCTGCTGAAGAGTA 1415
264 CTATGAGAAATTCATGACTCCCTCACAAGGAAAAGTGAGACAGATTTAGAGAAAGGTCTAGAGAC 323
1416 CTATGAGAAATGTTCCCTGCAAAAGCTGAGAGACCCAGAGAGTCTTTGGGAGGAATCTGAGAC 1475
324 AGAATATGCCCTTCTTAGGACTTCTGTTAGTAGGCCCTTG---TTCTGCACCCATGAGCA 380
1476 TGAGTATTCACCTTCTANATATGCTTCTACAGACCCAGGATGCCGATCCCCAGAGA 1535
381 TGATTATGAAGTTGTGTTTCCACAC 405
1536 TGAATATGAACCTTTTCATGCGCTCAC 1560

RESULT 5
ABX05108
ID ABX05108 standard; cDNA; 775 BP.
XX
XX ABX05108;
XX AC
XX AC
XX AC
DT 17-JAN-2003 (first entry)
DE
DE Human novel polynucleotide #123.
XX
XX Human; gene; ss; genetic disorder; gene mapping; medical imaging; cancer;
XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
XX Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
XX osteoarthritis; periodontal disease; liver fibrosis; viral infection;
XX fungal infection; bacterial infection; autoimmune disease; diabetes;
XX atopic dermatitis.
XX
XX Homo sapiens.
OS
XX
XX WO200274961-A1.
XX
XX
XX 26-SEP-2002.
XX
XX
XX 14-MAR-2002; 2002WO-US005109.
XX
XX 15-MAR-2001; 2001US-00810173.
XX
XX

PA (HYSE-) HYSQ INC.
XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
PI Wehrman T, Wang J, Wang D, Drmanac RT;
XX
DR WPI; 2003-040556/03.
DR P-PSDB; ABU00030.
XX
XX New isolated polypeptides and polynucleotides, useful for preventing,
PT treating or ameliorating medical conditions, such as cancer,
PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
PT disorders, and infections.
XX
XX Claim 1; SEQ ID NO 123; 235pp; English.
XX
XX The invention relates to human polynucleotides and the polypeptides they
CC encode. The polynucleotides and polypeptides are useful in diagnostics,
CC forensics, gene mapping, medical imaging, identification of mutations,
CC responsible for genetic disorders or other traits, assessing biodiversity
CC and producing many other types of data and products dependent on DNA and
CC amino acid sequences. They are also useful for preventing, treating or
CC ameliorating medical conditions, such as cancer, neurodegenerative
CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
CC invention. Note: The sequence data for this patent is not represented in
CC the printed specification but is based on sequence information supplied
CC by the European Patent Office
XX
SQ Sequence 775 BP; 226 A; 202 C; 160 G; 187 T; 0 U; 0 Other;

Query Match 11.5%; Score 46.6; DB 7; Length 775;
Best Local Similarity 50.8%; Pred. No. 0.0027;
Matches 165; Conservative 0; Mismatches 154; Indels 6; Gaps 2;

QY 87 ACGCAGGAATGACTACATCTTGAAGAAACCTTCAGATCAAGATGAAGAAAGCCCAAGA 146
DB 325 ATGGAAAAAATACTCATTTTGAAGAGGCGAAGATTCCTCAAAACGAAATGAAGAAT 384
QY 147 AGTTTCATCCACTTCTAATCAGGAAACAGAGAATGGCGATGGTTCTTGAAGAACTGTGCTA 206
DB 385 GTCATCTACTCCCATCCAGCAGGACATGTTGACCAGACCTACTCAGAGAGCTGTGCTA 444
QY 207 CACTGTCAATTAACA---CATCCCCCATCAGAGATCTCTCCCTGAGCTCCCAATCATATGG 263
DB 445 TACCTCATCAATCATCGGTTCTCTCAAGGCCCATCAGGGAATCTCTCTGAAGAGTA 504
QY 264 CTATGAGAACATTGACTCCCTCAAGGAAAGTGAGACAGTTTGAAGAAAGGTCAGAGAC 323
DB 505 CTATGAGAAATGTTCCCTGCAAGCTGAGAGACCCAGAGATCTCTGGGAGGAATCGAGAC 564
QY 324 AGAATATGCCCTTCTTAGGACTTCTGTGTAGAGCCCTTG---TTCCGTGACCCATGAGCA 380
DB 565 TGAGTATTAATCACTTCTACATATGCTTCTACAGACCCCGAGCATGCCCGATCCCGCAGAAGA 624
QY 381 TGATTATGAAGTCTGTGTTCCACAC 405
DB 625 TGAATATGAATCTTCTATGCTCTCAC 649

RESULT 6
ID AAS46328/c
XX AAS46328 standard; DNA; 6446 BP.
AC AAS46328;
XX
XX 18-DEC-2001 (first entry)
XX
XX Tumour suppressor gene derived chemically modified sequence #50.

KW Human; tumour suppressor gene; oncogene; antitumour; cytostatic; cancer;
KW tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;
XX cytosine methylation; ds.
OS Homo sapiens.
XX
XX WO200168912-A2.
XX
XX 20-SEP-2001.
XX
XX 15-MAR-2001; 2001WO-EP002955.
XX
XX 15-MAR-2000; 2000DE-01013847.
XX 06-APR-2000; 2000DE-01019058.
XX 07-APR-2000; 2000DE-01019173.
XX 30-JUN-2000; 2000DE-01032529.
XX 01-SEP-2000; 2000DE-01043826.
XX
XX (EPIG-) EPIGENOMICS AG.
XX
XX Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2001-602752/68.
XX
XX Fragments of chemically modified genes associated with tumor suppressor
XX genes and oncogenes useful in designing primers and probes for analyzing
XX diseases associated with cytosine methylation state e.g. cancer.
XX
XX Claim 1; SEQ ID NO 50; 27pp; English.
XX
XX The invention relates to a nucleic acid comprising a sequence of 18
XX bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with
XX bisulphite, of genes associated with tumour suppression and oncogenes
XX having a sequence taken from 536 (actually 533 since numbers 408, 458 and
XX 500 are missing from the sequence listing) sequences (Ss) and sequences
XX complementary to (Ss). The nucleic acid may be a peptide nucleic acid
XX oligomer (PNA) of at least 9 nucleotides and may form part of a set of
XX probes for detecting the cytosine methylation state and/or single
XX nucleotide polymorphisms and also to be used in an array for analysing
XX diseases associated with CpG dinucleotides e.g. cancers and tumours. The
XX probes can also be used in a method for ascertaining genetic and/or
XX epigenetic parameters for the diagnosis and/or therapy of existing
XX diseases or the predisposition to specific diseases, by analysing
XX cytosine methylations. The parameters may be compared to another set of
XX genetic and/or epigenetic parameters, the differences serving as basis
XX for diagnosis and/or prognosis events which are disadvantageous to
XX patients. The present sequence is one of the 533 genomic sequences
XX derived from tumour suppressor genes and oncogenes. Sequences with even
XX numbered Seq ID numbers are the complementary sequence of the
XX corresponding odd numbered sequence (e.g. ID 2 and ID1, ID 536 and ID
XX 535, except for those whose partner sequence is missing). Note: The
XX sequence data for this patent did not form part of the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 6446 BP; 1491 A; 208 C; 1796 G; 2951 T; 0 U; 0 Other;

Query Match 10.3%; Score 41.6; DB 4; Length 6446;
Best Local Similarity 56.6%; Pred. No. 0.16;
Matches 77; Conservative 0; Mismatches 59; Indels 0; Gaps 0;

QY 39 ACGGATCAAGAGAGCCCAAGAGAGGAAACCCAGATCAGGAAAGAAACGCGCAGGAAT 98
DB 1454 ACARAAAAAAGAAAAACCCCAACATAAACAACAAAAAAGAAAAACGCAAAAAA 1395
QY 99 GACTACATTTGAAGAAAACTTTCAAGATCAAGATGAAGAAAGCCAGAGTTTCATCCAC 158
DB 1394 AACGACGAAAAAAGAAAAATTTCTAAATTAAGAAAAAACAACAAAAAAGAAACGA 1335
QY 159 TTCTATCAGGAAC 174
DB 1334 TAAAAAAGCAAAAC 1319

Qy 39 AGAGAAATCAAAAAGAAGCCCAAGAAAGGAAACCCCGATGAGGAAAGAAAAACGGCAGGAAT 98
Db 18346 AAAATATTAAATAATTATATATATATATATATATATCTTTCACCAACCATATAACCCCACAAAAA 19387

Sequence	18988 BF	4826 A	358 C	1661 G	5244 T	0 G	0 C	0 G	0 C
Query Match	10.3%	Score 41.6	DB 6	Length 18988					
Best Local Similarity	47.3%	Score No. 0.24							
Matches 125	Conservative	0	Mismatches 139	Indels	0	Gaps	0		


```
XX SQ Sequence 1996 BP; 487 A; 410 C; 341 G; 758 T; 0 U; 0 Other;
Query Match 10.1%; Score 41; DB 4; Length 1996;
Best Local Similarity 48.5%; Pred. No. 0.16;
Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAATGACTACATTTGAAAGAAACTTC 121
DB 1187 ATGAAAGTTCAGAAAGAAACACAAACCCCATGAGTTCTCTGAAACGCTAAAGACC 1128
QY 122 AGATCAAGATAGAAAGCCAGAGTTTCATCCACTTCTATCAGGAAACGAGATG 181
DB 1127 AAGAGAAACTCAGAGTTATCTGAAACCTTGGAATTTACTATTCAGAGAGGAGATCA 1068
QY 182 GCAGTGGTCTGAGGAAGTGTGTACACTGTCTATTATCAGATCCCCCATCAGAGATCCT 241
DB 1067 AAGAGAGTCAAGAGAAAGAAAGACACTTTTAAAAAGATAGCCAAACAGATATTTCCCT 1008
QY 242 CCTGAGCTCCATGATGATGGCTATGAGACATTCAGTCCCTCCACAAAGGAA 294
DB 1007 CCAGCAAGGACCAAGAGAGTCTGAGTTGTTGATGACATACACTACTAGA 955
RESULT 13
ID ADC30201 standard; cDNA; 11924 BP.
XX AC ADC30201;
XX DT 18-DEC-2003 (first entry)
XX DE Human novel cDNA sequence, SEQ ID NO:283.
XX KW Human; diagnostic; drug screening; forensics; gene mapping;
XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX KW ulcers; osteoporosis; autoimmune disease; cancer;
XX KW molecular weight marker; food supplement; antiparkinsonian; neurotropic;
XX KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnary;
XX KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
XX KW gene therapy; chromosome 3; gene; ss.
XX OS Homo sapiens.
XX PN W02003029271-A2.
XX PD 10-APR-2003.
XX PF 24-SEP-2002; 2002WO-US030474.
XX PR 24-SEP-2001; 2001US-0324631P.
XX PA (HYSE-) HYSEQ INC.
XX PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
XX PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
XX PI Haley-Vicente D, Drmanac RT;
XX DR WPI; 2003-371981/35.
XX DR P-PSDB; ADC31172.
XX PT New polynucleotide and polypeptide useful for diagnosing, preventing or
XX PT treating conditions such as neurodegenerative diseases, anemias, platelet
XX PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
XX PT cancer.
XX PS Claim 1; SEQ ID NO 283; 1185pp; English.
XX CC The invention relates to 971 novel human cDNA sequences (ADC29919-
XX CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
XX CC invention also relates to nucleic acid sequences over 99% identical with
XX CC the novel human cDNAs. The invention additionally encompasses expression
```

```
CC vectors and host cells comprising a nucleic acid of the invention; the
CC recombinant production of a polypeptide of the invention; an antibody
CC against a polypeptide of the invention; a method of detecting
CC polynucleotides or polypeptides of the invention; and methods of
CC identifying a compound which binds to a polypeptide of the invention. The
CC invention further discloses methods of preventing, treating or
CC ameliorating a medical condition; kits comprising polynucleotide probes
CC and/or monoclonal antibodies for carrying out the methods of the
CC invention; methods for the identification of compounds that modulate the
CC expression or activity of the polynucleotide and/or polypeptide; and 767
CC contig sequences corresponding to the cDNA sequences of the invention
CC (ADC31961-ADC32627) and the polypeptides encoded by the contigs (ADC32628
CC -ADC33394). The nucleic acids and polypeptides of the invention are
CC useful in diagnostics, drug screening, forensics, gene mapping, in the
CC identification of mutations responsible for genetic disorders or other
CC traits, for assessing biodiversity, and in producing many other types of
CC data and products dependent on DNA and amino acid sequences. They are
CC also used for treating diseases such as Parkinson's disease, Alzheimer's
CC disease and other neurodegenerative diseases, anaemia, platelet
CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
CC cancer. The nucleic acids may also be used as hybridisation probes or
CC primers, and in the recombinant production of a protein. The polypeptides
CC are also useful in generating antibodies, as molecular weight markers,
CC and as food supplements. The present sequence represents a specifically
CC claimed human cDNA sequence of the invention. Note: The sequence data for
CC this patent did not form part of the printed specification, but was
CC obtained in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences.
```

```
XX SQ Sequence 11924 BP; 3959 A; 2706 C; 2427 G; 2832 T; 0 U; 0 Other;
Query Match 10.1%; Score 41; DB 9; Length 11924;
Best Local Similarity 48.5%; Pred. No. 0.3;
Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAATGACTACATTTGAAAGAAACTTC 121
DB 1048 ATGAAAGTTCAGAAAGAAAGAAACACAAACCCCATGAGTTCTCTGAAACGCTAAAGACC 1107
QY 122 AAGATCAAGATAGAAAGAAAGCCAGAGTTTCATCCACTTCTATCAGGAAACGAGATG 181
DB 1108 AAGAGAAACTCAGAGTTATCTGAAACCTTGGAATTTACTATTCAGAGAGGAGATCA 1167
QY 182 GCAGTGGTCTGAGGAAGTGTGTACACTGTCTATTATCAGATCCCCCATCAGAGATCCT 241
DB 1168 AAGAGAGTCAAGAGAAAGAAAGACACTTTTAAAAAGATAGCCAAACAGATATTTCCCT 1227
QY 242 CCTGAGCTCCATGATGATGGCTATGAGACATTCAGTCCCTCCACAAAGGAA 294
DB 1228 CCAGCAGGACCAAGAGAGTCTGAGTTGTTGATGACATACACTACTAGA 1280
```

```
RESULT 14
ID ADC30200 standard; cDNA; 11951 BP.
XX AC ADC30200;
XX DT 18-DEC-2003 (first entry)
XX DE Human novel cDNA sequence, SEQ ID NO:282.
XX KW Human; diagnostic; drug screening; forensics; gene mapping;
XX KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
XX KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
XX KW ulcers; osteoporosis; autoimmune disease; cancer;
XX KW molecular weight marker; food supplement; antiparkinsonian; neurotropic;
XX KW neuroprotective; antianaemic; anticoagulant; thrombolytic; vulnary;
XX KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
XX KW gene therapy; chromosome 3; gene; ss.
XX OS Homo sapiens.
```

PN WO2003029271-A2.
 XX 10-APR-2003.
 XX 24-SEP-2002; 2002WO-US030474.
 XX 24-SEP-2001; 2001US-0324631P.
 XX (HYSE-) HYSEQ INC.
 XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 XX Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 XX Haley-Vicente D, Drmanac RT;
 XX WPI; 2003-371981/35.
 XX P-PSDB; ADC31171.
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 XX treating conditions such as neurodegenerative diseases, anemias, platelet
 XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 XX cancer.
 XX Claim 1; SEQ ID NO 282; 1185pp; English.
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 XX ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 XX invention also relates to nucleic acid sequences over 99% identical with
 XX the novel human cDNAs. The invention additionally encompasses expression
 XX vectors and host cells comprising a nucleic acid of the invention; the
 XX recombinant production of a polypeptide of the invention; an antibody
 XX against a polypeptide of the invention; a method of detecting
 XX polynucleotides or polypeptides of the invention; and methods of
 XX identifying a compound which binds to a polypeptide of the invention. The
 XX invention further discloses methods of preventing, treating or
 XX ameliorating a medical condition; kits comprising polynucleotide probes
 XX and/or monoclonal antibodies for carrying out the methods of the
 XX invention; methods for the identification of compounds that modulate the
 XX expression or activity of the polynucleotide and/or polypeptide; and 767
 XX contig sequences corresponding to the cDNA sequences of the invention
 XX (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 XX -ADC33394). The nucleic acids and polypeptides of the invention are
 XX useful in diagnostics, drug screening, forensics, gene mapping, in the
 XX identification of mutations responsible for genetic disorders or other
 XX traits, for assessing biodiversity, and in producing many other types of
 XX data and products dependent on DNA and amino acid sequences. They are
 XX also used for treating diseases such as Parkinson's disease, Alzheimer's
 XX disease and other neurodegenerative diseases, anaemia, platelet
 XX disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 XX cancer. The nucleic acids may also be used as hybridisation probes or
 XX primers, and in the recombinant production of a protein. The polypeptides
 XX are also useful in generating antibodies, as molecular weight markers,
 XX and as food supplements. The present sequence represents a specifically
 XX claimed human cDNA sequence of the invention. Note: The sequence data for
 XX this patent did not form part of the printed specification, but was
 XX obtained in electronic format directly from WIPO at
 XX ftp.wipo.int/pub/published_pct_sequences.
 XX Sequence 11951 BP; 3967 A; 2711 C; 2434 G; 2839 T; 0 U; 0 Other;
 XX
 XX Query Match 10.1%; Score 41; DB 9; Length 11951;
 XX Best Local Similarity 48.5%; Pred. No. 0.3;
 XX Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
 QY 62 AAGGAACCCAGATCAGGAAGAAACGGCAGGAATGACTACATTGGAAGAAACTTC 121
 DB 1048 ATGAAGACTCAGAAAGAAACACACACCCATGAAGTTCTCTGACACGCTTAAGACC 1107
 QY 122 AAGATCAGATCAGAAAGCCAGGAAGTTTCATCCATCTTCAATCAGGAAGAAACGAGATG 181
 DB 1108 AAGAGAAACTCAGAGTTTATCTGAAACCTTGAAATTTACTATTTCAGAGAGAGATCA 1167
 QY 182 GCAGTGGTCTGAGAGAGTGCTACTGCTGCTAATTAATCATCCATCCCATCAGAGATCCT 241

Db 1168 AAGAGAGTCAAGAAAGAAAGGAAAGACACTTTTAAAAAGATAGCCACACAGATATTCCTT 1227
 QY 242 CCCTGAGCTCCATGATGATGCTATGAGACATGACTGCTCCTCACAAGGAAA 294
 Db 1228 CCAGCAGGACCATAAAGAGAGAGTCTGAGTTTGTGATGACATACTACTAGA 1280
 RESULT 15
 ACD19440
 ID ACD19440 standard; cDNA; 14994 BP.
 XX AC ACD19440;
 XX DT 25-AUG-2003 (first entry)
 XX DE CDNA encoding novel human protein #120.
 XX KW Human; NOV; gene therapy; endocrine related disease; diabetes;
 KW metabolism-related disease; obesity; central nervous system disorder;
 KW Alzheimer's disease; Parkinson's disease; epilepsy; multiple sclerosis;
 KW schizophrenia; depression; autoimmune disorder; inflammatory disorder;
 KW psoriasis; allergy; lupus erythematosus; asthma; cancer;
 KW inflammatory bowel disease; rheumatoid arthritis; osteoarthritis;
 KW colon cancer; lung cancer; liver cancer; breast cancer; ovarian cancer;
 KW prostate cancer; brain cancer; melanoma; liver disease; liver cirrhosis;
 KW lung disease; emphysema; obstructive pulmonary disease; haemophilia;
 KW stroke; infection; Gene; ss.
 XX KW Homo sapiens.
 OS
 XX WO2003023002-A2.
 XX PD 20-MAR-2003.
 XX PF 09-SEP-2002; 2002WO-US028539.
 XX PR 07-SEP-2001; 2001US-0318120P.
 XX PR 07-SEP-2001; 2001US-0318130P.
 XX PR 10-SEP-2001; 2001US-0318430P.
 XX PR 17-SEP-2001; 2001US-0322636P.
 XX PR 17-SEP-2001; 2001US-0322781P.
 XX PR 17-SEP-2001; 2001US-0322816P.
 XX PR 17-SEP-2001; 2001US-0322817P.
 XX PR 19-SEP-2001; 2001US-0323519P.
 XX PR 20-SEP-2001; 2001US-0323631P.
 XX PR 20-SEP-2001; 2001US-0323636P.
 XX PR 25-SEP-2001; 2001US-0324969P.
 XX PR 25-SEP-2001; 2001US-0325091P.
 XX PR 26-SEP-2001; 2001US-0324990P.
 XX PR 17-SEP-2002; 2002US-0373212P.
 XX PR 06-SEP-2002; 2002US-00236177.
 XX (CURA-) CURAGEN CORP.
 XX PA Spytek KA, Patturajan M, Gorman L, Li L, Anderson DW, Zhong M;
 XX PI Gerlach VL, Vernet CAM, Ellerman K, Berghs C, Rothenberg ME, Guo X;
 XX PI Shinkets RA, Leach MD, Catterton E, Kekuda R, Ji W, Miller CE;
 XX PI Rieger DK, Taupier RJ, Shenoy SG, Liu X, Padigar M, Aleobrook JP;
 XX PI Lepley DM, Edinger SR, Borgeas CE;
 XX WPI; 2003-313242/30.
 XX P-PSDB; ABO14747.
 XX New cytoplasmic, nuclear membrane bound or secreted polypeptides (NOVX)
 XX and polynucleotides, useful in gene therapy, e.g. for treating or
 XX preventing obesity, multiple sclerosis, allergy, cancers, hemophilia,
 XX stroke or infections.
 XX Claim 20; Page 336-340; 586pp; English.
 XX The invention describes a new isolated polypeptide (NOVX). The NOVX
 XX polypeptide, nucleic acid and antibody are useful as therapeutics,
 XX particularly in the manufacture of a medicament for treating a syndrome

CC associated with a human disease, which includes a pathology associated
CC with NOVX polypeptide. The DNA encoding the protein is useful in gene
CC therapy for treating the disease or condition. In particular, the NOVX
CC polypeptide or polynucleotide is useful for treating endocrine/
CC metabolism-related diseases (e.g. obesity or diabetes), central nervous
CC system disorders (e.g. Alzheimer's disease, Parkinson's disease,
CC epilepsy, multiple sclerosis, schizophrenia or depression), autoimmune
CC and inflammatory disorders (e.g. psoriasis, allergy, lupus erythematosus,
CC asthma, inflammatory bowel disease, rheumatoid arthritis or
CC osteoarthritis), cancers (e.g. colon, lung, liver, breast, ovarian,
CC prostate or brain cancers, or melanoma), liver diseases (e.g. liver
CC cirrhosis), lung diseases (emphysema or obstructive pulmonary disease),
CC haemophilia, stroke, or infections (e.g. viral, bacterial or parasitic).
CC These are also useful in developing powerful assay system for functional
CC analysis of various human disorders, as well as in diagnostic
CC applications, and for monitoring the effects of drugs during clinical
CC trials. This sequence encodes a novel human NOV protein
XX

SQ Sequence 14994 BP; 4902 A; 3673 C; 3081 G; 3338 T; 0 U; 0 Other;

Query Match 10.1%; Score 41; DB 7; Length 14994;

Best Local Similarity 48.5%; Pred. No. 0.33;

Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;

Qy 62 AAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACTTC 121

Db 4160 ATGAAAGTCAGAAAGAAACACACACCCCATGAGTTTCTCTGACACGCTTAAGACC 4219

Qy 122 AAGATCAAGATAAGAAAGCCAAAGAGTTTCATCCACTTCTATATCAGGAAACGAGATG 181

Db 4220 AAGAGAAACTCAGAGTTTATCTGAAACCTTGGAAATTACTATTTCAGAGGAGATCA 4279

Qy 182 GCAGTGGTTCTGAGAGTGCTGCTACACTGTCTATTATCATCCCCCATCAGAGATCCT 241

Db 4280 AAGAGATCAAGAGAGAGGAAAGACACTTTTAAAGATAGCCACAGATATTCCTT 4339

Qy 242 CCCTGAGCTCCATGATGATGGCTATGAGAACATTGACTCCCTCACAAGGAAA 294

Db 4340 CCAGCAGGACCATAAAGAGAGAGTCTGAGTTTGTGTGATGACATAACTACTAGA 4392

Search completed: April 22, 2004, 14:55:21

Job time : 359 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 15:25:28 ; Search time 336 Seconds
(without alignments)
5434.506 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405
Sequence: 1 atgggaattatctctgcg.....atgaagttgtttccacac 405

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 2907579 seqs, 2254313464 residues

Total number of hits satisfying chosen parameters: 5815158

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:*

- 1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
- 3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
- 4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
- 5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
- 6: /cgn2_6/ptodata/2/pubpna/PCFUS_PUBCOMB.seq:*
- 7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
- 8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
- 9: /cgn2_6/ptodata/2/pubpna/US09A_PUBCOMB.seq:*
- 10: /cgn2_6/ptodata/2/pubpna/US09B_PUBCOMB.seq:*
- 11: /cgn2_6/ptodata/2/pubpna/US09C_PUBCOMB.seq:*
- 12: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 13: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
- 14: /cgn2_6/ptodata/2/pubpna/US10A_PUBCOMB.seq:*
- 15: /cgn2_6/ptodata/2/pubpna/US10B_PUBCOMB.seq:*
- 16: /cgn2_6/ptodata/2/pubpna/US10C_PUBCOMB.seq:*
- 17: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
- 18: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
- 19: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	405	100.0	3762	15	US-10-005-907-1
C 2	41.6	10.3	6446	13	Sequence 1, Appl1
C 3	41.6	10.3	18988	13	Sequence 50, Appl1
C 4	41.6	10.3	18988	15	Sequence 63, Appl1
C 5	41.6	10.3	18988	15	Sequence 673, Appl1
C 6	41	10.1	1145	9	Sequence 61, Appl1
C 7	41	10.1	1196	9	Sequence 27195, A
C 8	40.4	10.0	1015	13	Sequence 10574, A
9	40	9.9	2921	9	Sequence 96139, A
10	39.8	9.8	785	9	Sequence 223, App
11	39.8	9.8	56258	13	Sequence 307, App
12	39.2	9.7	489	13	Sequence 213, App
13	39.2	9.7	963	13	Sequence 4490, App
14	39.2	9.7	966	13	Sequence 4492, App
					Sequence 4491, App

C 15	39.2	9.7	17869	15	US-10-311-455-78	Sequence 78, Appl1
C 16	39.2	9.7	17869	17	US-10-257-166-2	Sequence 2, Appl1
C 17	38.6	9.5	403	16	US-10-125-968-1156	Sequence 1156, Ap
C 18	38.4	9.5	347	13	US-10-424-599-101964	Sequence 101964,
C 19	38.4	9.5	737	13	US-10-027-632-157798	Sequence 157798,
C 20	38.4	9.5	737	16	US-10-027-632-157798	Sequence 157798,
C 21	38.4	9.5	6063	15	US-10-311-455-722	Sequence 722, App
C 22	38.2	9.4	683	13	US-10-424-599-30047	Sequence 30047, A
C 23	38.2	9.4	1937	13	US-10-282-122A-37450	Sequence 37450, A
C 24	38.2	9.4	7656	13	US-10-382-122A-37899	Sequence 37899, A
C 25	38	9.4	6112	15	US-10-311-455-636	Sequence 636, App
C 26	38	9.4	3673778	15	US-10-312-841-1	Sequence 1, Appl1
C 27	37.8	9.3	3673778	15	US-10-312-841-2	Sequence 2, Appl1
C 28	37.4	9.2	345	13	US-10-424-599-95566	Sequence 95566, A
C 29	37.4	9.2	554	9	US-09-878-178-2080	Sequence 2080, Ap
C 30	37.4	9.2	554	14	US-10-046-935-2080	Sequence 2080, Ap
C 31	37.4	9.2	554	15	US-10-146-502-2080	Sequence 1385, Ap
C 32	37.4	9.2	5437	15	US-10-311-455-1385	Sequence 87428, A
C 33	37.2	9.2	1896	13	US-10-424-599-87428	Sequence 1931, Ap
C 34	37.2	9.2	14006	15	US-10-311-455-1931	Sequence 18709, A
C 35	37	9.1	433	13	US-10-424-599-18709	Sequence 17883, A
C 36	37	9.1	544	10	US-09-814-353-17883	Sequence 71934, A
C 37	37	9.1	617	13	US-10-424-599-71934	Sequence 250669,
C 38	37	9.1	2131	13	US-10-027-632-250669	Sequence 250669,
C 39	37	9.1	2131	16	US-10-027-632-250669	Sequence 1610, Ap
C 40	37	9.1	7306	15	US-10-311-455-1610	Sequence 7092, Ap
C 41	37	9.1	14769	10	US-09-764-891-7092	Sequence 139, App
C 42	36.8	9.1	26392	13	US-10-087-192-139	Sequence 2228, Ap
C 43	36.6	9.0	408	9	US-09-938-842A-2228	Sequence 2228, Ap
C 44	36.6	9.0	408	11	US-09-938-842A-2228	Sequence 9958, Ap
C 45	36.6	9.0	451	10	US-09-918-995-9958	

ALIGNMENTS

RESULT 1

US-10-005-907-1
; Sequence 1, Application US/10005907
; Publication No. US20030166881A1
; GENERAL INFORMATION:
; APPLICANT: Union Chimique Belge, S.A.
; APPLICANT: No. US20030166881A1, Karl
; APPLICANT: Pirozzi, Gregory
; APPLICANT: Einstein, Richard
; TITLE OF INVENTION: NOVEL GENES ASSOCIATED WITH ALLERGIC HYPERSENSITIVITY AND MAST CELL
; FILE REFERENCE: 053529-5005
; CURRENT APPLICATION NUMBER: US/10/005,907
; CURRENT FILING DATE: 2001-12-07
; NUMBER OF SEQ ID NOS: 13
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 1
; LENGTH: 3762
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (25)..(432)
; OTHER INFORMATION:
US-10-005-907-1

Query Match 100.0%; Score 405; DB 15; Length 3762;
Best Local Similarity 100.0%; Pred. No. 1.1e-106; Mismatches 0; Indels 0; Gaps 0;
Matches 405; Conservative 0;
QY 1 ATGGGAATTTATCTCTCGAAACTCAGTTGCTGGAGAGAAATCAAAAGAGCCCAAG 60
25 ATGGGAATTTATCTCTCGAAACTCAGTTGCTGGAGAGAAATCAAAAGAGCCCAAG 84
Db 61 AAAGNAACCCAGTAGGAGAGAAAGCGGCGGAGAAATGACTTACATTTCGAAGAACTT 120
85 AAAGGAACCCAGTAGGAGAGAAAGCGGCGGAGAAATGACTTACATTTCGAAGAACTT 144


```

; PRIOR FILING DATE: 2001-07-02
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 2424
; SEQ ID NO 673
; LENGTH: 18988
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-311-455-673

Query Match          10.3%; Score 41.6; DB 15; Length 18988;
Best Local Similarity 47.3%; Pred. No. 0.52;
Matches 125; Conservative 0; Mismatches 139; Indels 0; Gaps 0;

QY 39 AGAAGATCAAAAGAACCCCAAGAAAGGAAACCCAGATGAGGAAGAAACGCGCAGGAAAT 98
DB 18346 AAAATATTAAATAATTATATCAATATATTAATCTATCAACACATATAACCCCAAAAAA 18397
QY 99 GACTACATTTGAAAGAAAACCTTCAAGATCAAGATAGAAAGAAACCCAGAGTTTCATCCAC 158
DB 18286 AACCAATTTCAAAAAAACCACCAAAATTCACCAACCTATATCAAAACACCATTTTTTCCTC 18227
QY 159 TTCTAATCAGGAAACCGAGATGCGCAGTGGTTCGTAAGAAGTGTCACACATGTCATTAA 218
DB 18226 TCTAATCATATAAACAACCTTTACCACTATATCATATTTATCATCACCCTACTATCAA 18167
QY 219 TCATATCCCCCATCAGAGATCTCCCTCGAGTCCCAATGATGATGCGTATGAGAAACATTGA 278
DB 18166 CAACCAATTTATCACAAAATATCTCTCTATCTTTCAAAACATCGAACCAACCTATCACCTTAA 18107
QY 279 CTCCTCTCACAAGGAAAGTGAGACA 302
DB 18106 ACCCTCTCCATAAAAAATACACA 18083

RESULT 5
US-10-240-485-61/c
; Sequence 61, Application US/10240485
; Publication No. US20030148327A1
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with
; TITLE OF INVENTION: Metastasis
; FILE REFERENCES: 5013.1007
; CURRENT APPLICATION NUMBER: US/10/240,485
; CURRENT FILING DATE: 2002-10-02
; PRIOR APPLICATION NUMBER: PCT/EP01/03970
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 202
; SEQ ID NO 61
; LENGTH: 18988
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-240-485-61

Query Match          10.3%; Score 41.6; DB 15; Length 18988;
Best Local Similarity 47.3%; Pred. No. 0.52;

```

NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 27195
LENGTH: 1145
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004886.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
OTHER INFORMATION: EST HUMAN HIT: A1636014.1, EVALUE 0.00e+00
OTHER INFORMATION: SWISSPROT HIT: P38254, EVALUE 7.50e-01
OTHER INFORMATION: NT HIT: AF138789.2, EVALUE 0.00e+00
US-09-864-761-27195

Query Match 10.1%; Score 41; DB 9; Length 1145;
Best Local Similarity 48.5%; Pred. No. 0.21;
Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 62 AAGGAACCCAGATGAGGAAGAAACGGCGGGAATGACTACATTTGGAAGAAACTTC 121
DB 877 ATGAAAAGTTCAGAAAAGAAACACACACCCCATGAAAGTTCTCTGAAACGCCCTAAAGACC 818
QY 122 AAGATCAAGATAAGAAAAGCCCAAGAGTTTTCATCCACTTCTAATCAGGAAAACGAGAAATG 181
DB 817 AAGAAAACCTCAGAGTTATCTGAACCTTGGAAATTAATTTTCAAGAGAGGATCA 758
QY 182 GCAGTGGTTTGAAGAGTGTGCTACTGTCTATTAATCAATCATCCCCCATCAGAGATCCT 241
DB 757 AAGAGAGTCAAGAAAAGAAAGAACACACTTTTAAAAAAGATAGCAACAAGATATTCCTT 698
QY 242 CCCTGAGCTCCAATGATGCTGCTATGAGAACATTTGACTCCCTCCACAAAGGAAA 294
DB 697 CCAGCAGGACCAATAAGAGAAGTCTGAGTTTGTGTGATGACATACTACTAGA 645

RESULT 7
US-09-864-761-10574/c
Sequence 10574, Application US/09864761
Patent No. US20020048763A1
GENERAL INFORMATION:
APPLICANT: Penn, Sharon G.
APPLICANT: Rank, David R.
APPLICANT: Hanzel, David K.
APPLICANT: Chen, Wensheng
TITLE OF INVENTION: HUMAN GENOME-DERIVED SINGLE EXON NUCLEIC ACID PROBES USEFUL FOR
TITLE OF INVENTION: GENE EXPRESSION ANALYSIS BY MICROARRAY
FILE REFERENCE: Acomica-X-1
CURRENT APPLICATION NUMBER: US/09/864,761
PRIOR FILING DATE: 2001-05-23
PRIOR APPLICATION NUMBER: US 60/180,312
PRIOR FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: US 60/207,456
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: US 09/632,366
PRIOR FILING DATE: 2000-08-03
PRIOR APPLICATION NUMBER: GB 24263.6
PRIOR FILING DATE: 2000-10-04
PRIOR APPLICATION NUMBER: US 60/236,359
PRIOR FILING DATE: 2000-09-27
PRIOR APPLICATION NUMBER: PCT/US01/00666
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00667
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00664
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00669
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00665
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00668
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00663
PRIOR FILING DATE: 2001-01-30

PRIOR APPLICATION NUMBER: PCT/US01/00662
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00661
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: PCT/US01/00670
PRIOR FILING DATE: 2001-01-30
PRIOR APPLICATION NUMBER: US 60/234,687
PRIOR FILING DATE: 2000-09-21
PRIOR APPLICATION NUMBER: US 09/608,408
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: US 09/774,203
PRIOR FILING DATE: 2001-01-29
NUMBER OF SEQ ID NOS: 49117
SOFTWARE: Annonax Sequence Listing Engine vers. 1.1
SEQ ID NO 10574
LENGTH: 1996
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
OTHER INFORMATION: MAP TO AC004886.1
OTHER INFORMATION: EXPRESSED IN BRAIN, SIGNAL = 1.1
US-09-864-761-10574

Query Match 10.1%; Score 41; DB 9; Length 1996;
Best Local Similarity 48.5%; Pred. No. 0.27;
Matches 113; Conservative 0; Mismatches 120; Indels 0; Gaps 0;
QY 62 AAGGAACCCAGATGAGGAAGAAACGGCGGGAATGACTACATTTGGAAGAAACTTC 121
DB 1187 ATGAAAAGTTCAGAAAAGAAACACACACCCCATGAAAGTTTCTCTGAAACGCCCTAAAGACC 1128
QY 122 AAGATCAAGATAAGAAAAGCCCAAGAGTTTTCATCCACTTCTAATCAGGAAAACGAGAAATG 181
DB 1127 AAGAAAACCTCAGAGTTATCTGAACCTTGGAAATTAATTTTCAAGAGAGGATCA 1068
QY 182 GCAGTGGTTTGAAGAGTGTGCTACTGTCTAATTAATCAATCCCCCATCAGAGATCCT 241
DB 1067 AAGAGAGTCAAGAAAAGAAAGAACACTTTTAAAAAGATAGCAACAAGATATTCCTT 1008
QY 242 CCCTGAGCTCCAATGATGCTGCTATGAGAACATTTGACTCCCTCCACAAAGGAAA 294
DB 1007 CCAGCAGGACCAATAAGAGAAGTCTGAGTTTGTGTGATGACATACTACTAGA 955

RESULT 8
US-10-424-599-96139
Sequence 96139, Application US/10424599
Publication No. US20040031072A1
GENERAL INFORMATION:
APPLICANT: La Rosa Thomas J
APPLICANT: Kovalic David K
APPLICANT: Zhou Yihua
APPLICANT: Cao Yongwei
TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
FILE REFERENCE: 38-21(53223)B
CURRENT APPLICATION NUMBER: US/10/424,599
CURRENT FILING DATE: 2003-04-28
NUMBER OF SEQ ID NOS: 285684
SEQ ID NO 96139
LENGTH: 1015
TYPE: DNA
ORGANISM: Glycine max
FEATURE:
OTHER INFORMATION: Clone ID: PAT_MBT3847_57824C.1
US-10-424-599-96139

Query Match 10.0%; Score 40.4; DB 13; Length 1015;
Best Local Similarity 54.8%; Pred. No. 0.3;
Matches 80; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 2 TGGGAAATTATCTCTCGGAAACTCAGTGTGCTGGGAGAGATCAAAAGAGCCCAAGA 61

Db 735 TGGATGATTATATAGAGAAATACTCAGTTCTTCTGCTAAAAAAGAAAAA 794
Qy 62 AAGGAACCCAGATGAGGAAGAAACCGCAGCAATGACTACATTGGAAGAACTTC 121
Db 795 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 854
Qy 122 AAGATCAAGATAGAAAGAAAGCCAGAA 147
Db 855 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 880

RESULT 9

US-09-925-301-223
; Sequence 223, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 223
; LENGTH: 2921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1609)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2919)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2920)
; OTHER INFORMATION: n equals a,t,g, or c
; NAME/KEY: misc feature
; LOCATION: (2921)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-925-301-223

Query Match 9.9%; Score 40; DB 9; Length 2921;
Best Local Similarity 54.1%; Pred. No. 0.63;
Matches 79; Conservative 1; Mismatches 66; Indels 0; Gaps 0;
Qy 2 TGGGAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGA 61
Db 2742 TGGGAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGA 2801
Qy 62 AAGGAACCCAGATGAGGAAGAAACCGCAGCAATGACTACATTGGAAGAACTTC 121
Db 2802 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2861
Qy 122 AAGATCAAGATAGAAAGAAAGCCAGAA 147
Db 2862 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 2887

RESULT 10

US-09-925-301-307
; Sequence 307, Application US/09925301
; Patent No. US20020052308A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins and Antibodies
; FILE REFERENCE: PA106
; CURRENT APPLICATION NUMBER: US/09/925,301
; CURRENT FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: PCT/US00/05882
; PRIOR FILING DATE: 2000-03-08
; PRIOR APPLICATION NUMBER: 60/124,270
; PRIOR FILING DATE: 1999-03-12
; NUMBER OF SEQ ID NOS: 1694
; SOFTWARE: Patent in Ver. 2.0
; SEQ ID NO 307
; LENGTH: 785
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-925-301-307

Query Match 9.8%; Score 39.8; DB 9; Length 785;
Best Local Similarity 54.4%; Pred. No. 0.39;
Matches 80; Conservative 0; Mismatches 67; Indels 0; Gaps 0;
Qy 1 ATGGGAATTATCTCTCGGAAATCTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAG 60
Db 629 ATTGGTATCTCCTATACAGAAAGCCCAATAGACCCAAAGTCTCCAAAGAGCTCAAG 688
Qy 61 AAAGGAACCCAGATGAGGAAGAAACCGCAGCAATGACTACATTGGAAGAACTTC 120
Db 689 GAAGAGAAACGAAACAGAGCAAAAGAAATATAATAATAATAATAATAATAATAATA 748
Qy 121 CAAGATCAAGATAGAAAGAAAGCCAGAA 147
Db 749 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 775

RESULT 11

US-10-087-192-913/c
; Sequence 913, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 913
; LENGTH: 56258
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)....(56258)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-913

Query Match 9.8%; Score 39.8; DB 13; Length 56258;
Best Local Similarity 60.7%; Pred. No. 2.8;
Matches 65; Conservative 0; Mismatches 42; Indels 0; Gaps 0;
Qy 40 GGAATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAGAAACCGCAGCAATG 99
Db 11067 GAGAAGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 11008
Qy 100 ACTACATTTTGAAGAAATCTCAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 146
Db 11007 AAAGGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 10961

RESULT 12

US-10-335-977-4490
; Sequence 4490, Application US/10335977

```
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 4490:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 489 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...489
; SEQUENCE DESCRIPTION: SEQ ID NO: 4490:
US-10-335-977-4490
Query Match 9.7%; Score 39.2; DB 13; Length 489;
Best Local Similarity 64.1%; Pred. No. 0.47;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 57 CAAGAAAGGAAACCCAGATGCGAAGAAACGGCAGGAATGACTACATTGGAAGAAA 116
Db 108 CAATGAACCTAATCTCTACAGAAAGAAAAGTTAGACCAAGCCTAAATTAGAAGAAA 167

QY 117 ACTTCAAGATCAAGATAAGAAAGCCAAAG 148
Db 168 ACACAAAGAACCAAGCAAGCAAGCAAG 199

RESULT 13
US-10-335-977-4492
; Sequence 4492, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
```

```
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: IBM PC Compatible
; OPERATING SYSTEM: Windows NT 4.0
; SOFTWARE: UNIX
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/335,977
; FILING DATE: 30-Dec-2002
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/993,002
; FILING DATE: 17-DEC-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Mandragouras, Amy E.
; REGISTRATION NUMBER: 36,207
; REFERENCE/DOCKET NUMBER: GTN-018
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 227-7400
; TELEFAX: (617) 742-4214
; INFORMATION FOR SEQ ID NO: 4492:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 963 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: circular
; MOLECULE TYPE: DNA (genomic)
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Helicobacter pylori
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...963
; SEQUENCE DESCRIPTION: SEQ ID NO: 4492:
US-10-335-977-4492
Query Match 9.7%; Score 39.2; DB 13; Length 963;
Best Local Similarity 64.1%; Pred. No. 0.65;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;

QY 57 CAAGAAAGGAAACCCAGATGCGAAGAAACGGCAGGAATGACTACATTGGAAGAAA 116
Db 582 CAATGAACCTAATCTCTACAGAAAGAAAAGTTAGACCAAGCCTAAATTAGAAGAAA 641

QY 117 ACTTCAAGATCAAGATAAGAAAGCCAAAG 148
Db 642 ACACAAAGAACCAAGCAAGCAAGCAAG 673

RESULT 14
US-10-335-977-4491
; Sequence 4491, Application US/10335977
; Publication No. US20040052799A1
; GENERAL INFORMATION:
; APPLICANT: DOUGLAS SMITH et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES
; RELATING TO HELICOBACTER PYLORI FOR
; DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 10031
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: LAHIVE & COCKFIELD
; STREET: 28 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109-1875
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
```

COMPUTER: IBM PC Compatible
OPERATING SYSTEM: Windows NT 4.0
SOFTWARE: UNIX
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/10/335,977
FILING DATE: 30-Dec-2002
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/993,002
FILING DATE: 17-DEC-1997
ATTORNEY/AGENT INFORMATION:
NAME: Mandragoras, Amy E.
REGISTRATION NUMBER: 36,207
REFERENCE/DOCKET NUMBER: GYN-018
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617)227-7400
TELEFAX: (617)742-4214
INFORMATION FOR SEQ ID NO: 4491:
SEQUENCE CHARACTERISTICS:
LENGTH: 966 base pairs
TYPE: nucleic acid
STRADEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Helicobacter pylori
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...966
SEQUENCE DESCRIPTION: SEQ ID NO: 4491:
US-10-335-977-4491

Query Match 9.7%; Score 39.2; DB 13; Length 966;
Best Local Similarity 64.1%; Pred. No. 0.65;
Matches 59; Conservative 0; Mismatches 33; Indels 0; Gaps 0;
Qy 57 CAGGAAAGGAAACCCAGATGAGGAAAGAAACGCGCAGGAAATGACTACATTTCGAAAGAA 116
Db 582 CAATGAACTATCTTACAG 641
Qy 117 ACTTCAAGATCAAGATAGAAAGAAAGCCCAAGAG 148
Db 642 ACACAAAGAACACAGACAGCAAGCAAGCAAGCAAGAG 673

RESULT 15
US-10-311-455-78/c
Sequence 78, Application US/10311455
Publication No. US20030143606A1
GENERAL INFORMATION:
APPLICANT: OLEK, Alexander
APPLICANT: PIEPENBROCK, Christian
APPLICANT: BERLIN, Kurt
TITLE OF INVENTION: Diagnosis of Diseases Associated with the Immune System by Determining Cytosine Methylation
FILE REFERENCE: 5013.1014
CURRENT APPLICATION NUMBER: US/10/311.455
CURRENT FILING DATE: 2002-12-16
PRIOR APPLICATION NUMBER: PCT/EP01/07537
PRIOR FILING DATE: 2001-07-02
PRIOR APPLICATION NUMBER: DE 10032529.7
PRIOR FILING DATE: 2000-06-30
PRIOR APPLICATION NUMBER: DE 10043826.1
PRIOR FILING DATE: 2000-09-01
NUMBER OF SEQ ID NOS: 2424
SEQ ID NO 78
LENGTH: 17869
TYPE: DNA
ORGANISM: Artificial Sequence
FEATURE:
OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
FEATURE:

NAME/KEY: unsure
LOCATION: 3465, 3586
OTHER INFORMATION: n is a or g or c or t
US-10-311-455-78
Query Match 9.7%; Score 39.2; DB 15; Length 17869;
Best Local Similarity 47.9%; Pred. No. 2.5;
Matches 113; Conservative 0; Mismatches 123; Indels 0; Gaps 0;
Qy 6 AAATTATCTCTCGGAAACATCAGTTGCTGGGAGAGATCAAAAGAGAGCCCAAGAAAGG 65
Db 8214 AAATATCCAAAACTATCTTAATAATAATAATAATAATAATAATAATAATAATAATA 8155
Qy 66 AAACCCAGATGAGGAAAGAGAAACCGCAGGAAATGACTACATTTGAAAGAAACCTTCAAGA 125
Db 8154 AAACCTAAATACAAACTTAAACCAAAATCGACTAACGAATTTTAAAAAAATCACTAAA 8095
Qy 126 TCAGATTAAGAAAGCCCAAGAGTTTCATCTTCAATCAGGAAACGAGAGATGGCAG 185
Db 8094 AATCTTAAACAAAAAAAACTTATATTCACGCAAACTAATTATATATATATATAT 8035
Qy 186 TGGTTCTGAAGAGTGTGCTACACTGTGCTATTAATCAATCCCTCATCAGAGATCCT 241
Db 8034 TTTCACTTAAAAAATATTTTAACTTCTTCCAACTAAACAAATATTTTATTACT 7979

Search completed: April 22, 2004, 17:14:09
Job time : 347 secs

Result No.	Score	Query		DB	ID	Description
		Match	Length			
C 1	49	12.1	7218	1	US-08-322-463-14	Sequence 14, Appl
C 2	38.2	9.4	1092	4	US-09-328-352-1567	Sequence 1567, Ap
C 3	36.8	9.1	549	4	US-09-107-532A-3987	Sequence 2987, Ap
C 4	36.4	9.0	580073	4	US-08-545-528D-1	Sequence 1, Appl
C 5	35.4	8.7	2608	4	US-09-304-615-16	Sequence 16, Appl
C 6	35	8.6	759	4	US-09-540-236-1733	Sequence 1733, Ap
C 7	35	8.6	2447	2	US-09-014-969-14	Sequence 14, Appl
C 8	34.8	8.6	339	4	US-09-621-976-8976	Sequence 8976, Ap
C 9	34.8	8.6	118067	4	US-09-134-001C-1161	Sequence 1161, Ap
C 10	34.8	8.6	118067	4	US-09-497-855A-32	Sequence 32, Appl
C 11	34	8.4	1462	4	US-09-133-962A-5	Sequence 5, Appl
C 12	34	8.4	1985	4	US-09-907-794A-212	Sequence 212, App
C 13	34	8.4	1985	4	US-09-905-125A-212	Sequence 212, App
C 14	34	8.4	1985	4	US-09-902-775A-212	Sequence 212, App
C 15	33.8	8.3	5340	4	US-09-627-123-21	Sequence 21, Appl
C 16	33.6	8.3	249	4	US-09-621-976-1322	Sequence 1322, Ap
C 17	33.6	8.3	3275	4	US-09-370-838-151	Sequence 151, App
C 18	33.4	8.2	240	4	US-09-621-976-1324	Sequence 1324, Ap
C 19	33.4	8.2	640	4	US-09-976-594-1142	Sequence 1142, Ap
C 20	33.2	8.2	926	3	US-08-945-994-8	Sequence 8, Appl
C 21	33.2	8.2	5727	4	US-08-956-171E-190	Sequence 190, App
C 22	33	8.1	400	4	US-08-956-171E-4053	Sequence 4053, Ap
C 23	33	8.1	674	4	US-09-620-405B-465	Sequence 465, App
C 24	33	8.1	674	4	US-09-433-828B-465	Sequence 465, App
C 25	33	8.1	674	4	US-09-604-287A-465	Sequence 465, App
C 26	33	8.1	674	4	US-09-634-759-465	Sequence 465, App
C 27	33	8.1	1000	3	US-09-018-584A-34	Sequence 34, Appl

Best Local Similarity 2.2%; Pred. No. 4e-05;
Matches 7; Conservative 192; Mismatches 122; Indels 0; Gaps 0;
QY 7 AATTATCTCTCGGAAACTCAGTTGCTGGAGAGATCAAAAGAGCCCAAGAAAGGA 66
Db 1447 AAGAAATTTGGTACRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRRR 1388
QY 67 AACCCAGATGAGGAAAGAAACGGCAGGAATGACTATCTTGAAGAAACTTCAAGAT 126
Db 1387 RRR 1328
QY 127 CAAGTAAAGAAAGCAAGAGTTTCATCCACTTCTAATCAGAAACGAGATGGCAGT 186
Db 1327 RRR 1268
QY 187 GCTTCTGAAGAGTGTGTACACTGTCTATTATCATCATCCCATCAGAGATCTCCCTG 246
Db 1267 RRR 1208
QY 247 AGCTCAATGATGCTGTATGAGAACTTCACTCCCTCAAGAAAGTGAAGACATTT 306
Db 1207 RRR 1148
QY 307 AGAGAAAGTGTACAGACAGAA 327
Db 1147 RRRRRRRRRRRRRRRRRRR 1127

RESULT 2

US-09-328-352-1567/c
; Sequence 1567, Application US/09328352
; Patent No. 6562958
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER
; FILE REFERENCE: GTC99-03PA
; CURRENT APPLICATION NUMBER: US/09/328,352
; CURRENT FILING DATE: 1999-06-04
; NUMBER OF SEQ ID NOS: 8252
; SEQ ID NO 1567
; LENGTH: 1092
; TYPE: DNA
; ORGANISM: Acinetobacter baumannii
US-09-328-352-1567

Query Match 9.4%; Score 38.2; DB 4; Length 1092;
Best Local Similarity 52.9%; Pred. No. 0.037; 73; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 73; Indels 0; Gaps 0;
QY 43 AATCAAAAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACT 102
Db 315 AATAATTAGAAACCAATAAACCCAGGTTTCAGAGGTGATAACCCACGGGTGTAATGGGT 256
QY 103 ACATTTGAAGAAACTTCAAGATCAAGTAAAGAAACCCAGAGTTTCATCCACTTCT 162
Db 255 ACGGTAGAAAGACAATACTTGGGGCAATGATATGTTTCAAGAACCATGCCCCAATACT 196
QY 163 AATCAGAAACGAGATGGCAGTGGTTCTGAAGA 197
Db 195 GCTCATTAAGAAAGAAATTTGGCTGCAAGACAGACA 161

RESULT 3

US-09-107-532A-2987
; Sequence 2987, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:

ADDRESSEE: GENOME THERAPEUTICS CORPORATION
STREET: 100 Beaver Street
CITY: Waltham
STATE: Massachusetts
COUNTRY: USA
ZIP: 02354
COMPUTER READABLE FORM:
MEDIUM TYPE: CD-ROM ISO9660
COMPUTER: PC
OPERATING SYSTEM: <Unknown>
SOFTWARE: ASCII
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/107,532A
FILING DATE: 30-Jun-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/085,598
FILING DATE: 14 May 1998
APPLICATION NUMBER: 60/051571
FILING DATE: July 2, 1997
ATTORNEY/AGENT INFORMATION:
NAME: Ariniello, Pamela Deneske
REGISTRATION NUMBER: 40,489
REFERENCE/DOCKET NUMBER: GTC-012
TELECOMMUNICATION INFORMATION:
TELEPHONE: (781)893-5007
TELEFAX: (781)893-8277
INFORMATION FOR SEQ ID NO: 2987:
SEQUENCE CHARACTERISTICS:
LENGTH: 549 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: circular
MOLECULE TYPE: DNA (genomic)
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Enterococcus faecium
FEATURE:
NAME/KEY: misc feature
LOCATION: (B) LOCATION 1...549
SEQUENCE DESCRIPTION: SEQ ID NO: 2987:
US-09-107-532A-2987

Query Match 9.1%; Score 36.8; DB 4; Length 549;
Best Local Similarity 50.0%; Pred. No. 0.072;
Matches 92; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
QY 35 TGGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAAAGAAACCGCAGG 94
Db 29 TTGGGAACAATAAGAGACTCGCCAAATAAAGAAAGAGCAGCAAGAAAGAAAGCAGC 88
QY 95 AATGACTACTATTGAAAGAAACTTCAAGATCAAGATAAGAAAGCCCAAGAGTTTCTAT 154
Db 89 AAGCAATTACCATCTGAAATAAGCAATTAGAAAGTGCCACCAAGTTCAAGAAATAAAAA 148
QY 155 CACATTCATCAGGAAACGAGATGGCAGTGGTTCTGAGAAAGTGTGCTACACTGCA 214
Db 149 CTTCTCCCAATTTGAAAGACAGAGTAGAGGAAGAACAAAGAAAGTTGGCCCAATCA 208
QY 215 TTAA 218
Db 209 GAA 212

RESULT 4

US-08-545-528D-1/c
; Sequence 1, Application US/08545528D
; Patent No. 6537773
; GENERAL INFORMATION:
; APPLICANT: Fraser et al.
; TITLE OF INVENTION: Nucleotide Sequence of the Mycoplasma Genitium Genome, Fragments
; Patent No. 6537773
; TITLE OF INVENTION: Thereof, and Uses Thereof

```

Db      291  TGCCTAATGCATAGAACTTAAAGCAGCTTAAATAGGCTCAATGGTCAGCG
Qy      339  TAGGACTTCTGTAGTAGGCCTT 361
Db      351  AAGCTATCTTTATATCACCCCTT 373

RESULT 7
US-09-014-969-14
; Sequence 14, Application US/09014969
; Patent No. 5965397
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: LaVallie, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 CambridgePark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/014,969
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Sprunger, Suzanne A.

```


REGISTRATION NUMBER: 41,323
TELECOMMUNICATION INFORMATION:
TELEPHONE: (617) 498-8284
TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 14:
SEQUENCE CHARACTERISTICS:
LENGTH: 2447 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
MOLECULE TYPE: cDNA
US-09-014-969-14

Query Match 8.6%; Score 35; DB 2; Length 2447;
Best Local Similarity 50.9%; Pred. No. 0.52; Mismatches 80; Indels 0; Gaps 0;
Matches 83; Conservative 0;
QY 17 TCGAAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGAAAGGAAACCCAGATG 76
Db 2221 TGTGCAATATAACCTTTTGGAGAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 2280
QY 77 AGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACCTTCAAGATCAAGATGAAGA 136
Db 2281 AAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 2340
QY 137 AAGCCAGAGATTTCATCCACTTCTAATCAGGAAACGAGAA 179
Db 2341 AAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAAARAAAAA 2383

RESULT 8
US-09-621-976-8976
Sequence 8976, Application US/09621976
Patent No. 6639063
GENERAL INFORMATION:
APPLICANT: Dumas Milne Edwards, J.B.
APPLICANT: Jobert, S.
APPLICANT: Giordano, J.Y.
TITLE OF INVENTION: ESTs and Encoded Human Proteins.
FILE REFERENCE: GENSET.054PR2
CURRENT APPLICATION NUMBER: US/09/621,976
CURRENT FILING DATE: 2000-07-21
NUMBER OF SEQ ID NOS: 19335
SOFTWARE: Patent.pm
SEQ ID NO 8976
LENGTH: 399
TYPE: DNA
ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match 8.6%; Score 34.8; DB 4; Length 399;
Best Local Similarity 12.4%; Pred. No. 0.26; Mismatches 110; Indels 0; Gaps 0;
Matches 30; Conservative 110;
QY 3 GCGAATATCTCTCCGCAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGAA 62
Db 42 KRRRRRAAWKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSKMSK 101
QY 63 AGAAACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAACCTTCA 122
Db 102 YRYRKACTKWRAGCMGKAGWAWAYAKWYMWAWRRTAMKYWAMKSKMSMRERRAWYA 161
QY 123 AGATCAAGATGAAGAAACCGCAAGTTCATCCACTTCTAATCAGGAAACGAGAAATGG 182
Db 162 WYMYWARTWGWPCASCYRGYMASAGYMYWYMYWYMYWYMYWYMYWYMYWYMYWYMYW 221
QY 183 CAGTGGTTCTGAGAGAGTGTGTACACTGTCTAATTAATCAATCCCATCCATCAGATCCTC 242
Db 222 YSCWSSYCWGAKMMWYKTSRWSYWSYRCKYRRCSCCWSSNSCYWKTYRWSYWCASC 281
QY 243 CC 244
Db 282 Y 283

RESULT 9
US-09-134-001C-1161
Sequence 1161, Application US/09134001C
Patent No. 6380370
GENERAL INFORMATION:
APPLICANT: Lynn Doucette-Stamm et al
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS
TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: GTC-007
CURRENT APPLICATION NUMBER: US/09/134,001C
CURRENT FILING DATE: 1998-08-13
PRIOR APPLICATION NUMBER: US 60/064,964
PRIOR FILING DATE: 1997-11-08
PRIOR APPLICATION NUMBER: US 60/055,779
PRIOR FILING DATE: 1997-08-14
NUMBER OF SEQ ID NOS: 5674
SEQ ID NO 1161
LENGTH: 729
TYPE: DNA
ORGANISM: Staphylococcus epidermidis
US-09-134-001C-1161

Query Match 8.6%; Score 34.8; DB 4; Length 729;
Best Local Similarity 49.5%; Pred. No. 0.34; Mismatches 90; Conservative 0; Indels 0; Gaps 0;
Matches 90;
QY 40 GAGATCAAAAGAGCCCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 99
Db 358 GAACACGTTAAACAAAGAAAGAAACCATCAAGATAAACAAAGACGTCGCGATAAG 417
QY 100 ACTACATTTGAAAGAAACCTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAG 159
Db 418 CCATCAATGTGAGAAATATTATAGTGAAGTTTGACGATTTAGAGCAAGTTATCGACTAT 477
QY 160 TCTATCAGAAACGAGAAATGGCAGTGGTCTTGAAGAGTGTGCTACACTGTCATTAAT 219
Db 478 GCTTATCATTAATCAAAACACTGATGAATTTGAAGATTTATTATATATATGATGATAAT 537
QY 220 CA 221
Db 538 AA 539

RESULT 10
US-09-497-855A-32/c
Sequence 32, Application US/09497855A
Patent No. 6605432
GENERAL INFORMATION:
APPLICANT: Huang, Tim
TITLE OF INVENTION: HIGH-THROUGHPUT METHODS FOR DETECTING DNA METHYLATION
FILE REFERENCE: UMO1523
CURRENT APPLICATION NUMBER: US/09/497,855A
CURRENT FILING DATE: 2000-02-04
PRIOR APPLICATION NUMBER: 60/120,592
PRIOR FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: 60/118,760
PRIOR FILING DATE: 1999-02-05
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PatentIn version 3.0
SEQ ID NO 32
LENGTH: 118067
TYPE: DNA
ORGANISM: Homo sapiens;
US-09-497-855A-32

Query Match 8.6%; Score 34.8; DB 4; Length 118067;
Best Local Similarity 53.7%; Pred. No. 3.5; Mismatches 72; Conservative 0; Indels 0; Gaps 0;
Matches 72;
QY 37 GGAGAGATCAAAAGAGCCCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 96
Db 37 GGAGAGATCAAAAGAGCCCAAGAAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 96

; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-907-794A-212

Query Match 8.4%; Score 34; DB 4; Length 1985;
Best Local Similarity 52.1%; Pred. No. 0.95;
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 2 TGGGAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGA 61
Db 1824 TTGGACATAAATGGTGTATGCTTCCGCCAAAAAAGAAAAA 1883

QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTACATTTCAAAAGAACTTC 121
Db 1894 AA 1943

QY 122 AAGATCAAGATAAGAAAGCCCAAGA 147
Db 1944 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1969

RESULT 13

US-09-905-125A-212
; Sequence 212, Application US/0905125A
; Patent No. 6664376
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gernitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.
; APPLICANT: Kijavlin, Ivar J.
; APPLICANT: Mather, Jennie P.
; APPLICANT: Pan, James
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William, I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids encoding the Same
; FILE REFERENCE: 10466-14
; CURRENT APPLICATION NUMBER: US/09/905,125A
; CURRENT FILING DATE: 2001-07-12
; PRIOR APPLICATION NUMBER: PCT/US00/04414
; PRIOR FILING DATE: 2000-02-22
; PRIOR APPLICATION NUMBER: US 60/143,048
; PRIOR FILING DATE: 1999-07-07
; PRIOR APPLICATION NUMBER: US 60/145,698
; PRIOR FILING DATE: 1999-07-26

; PRIOR APPLICATION NUMBER: US 60/146,222
; PRIOR FILING DATE: 1999-07-28
; PRIOR APPLICATION NUMBER: PCT/US99/20594
; PRIOR FILING DATE: 1999-09-08
; PRIOR APPLICATION NUMBER: PCT/US99/20944
; PRIOR FILING DATE: 1999-09-13
; PRIOR APPLICATION NUMBER: PCT/US99/21090
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/21547
; PRIOR FILING DATE: 1999-09-15
; PRIOR APPLICATION NUMBER: PCT/US99/23089
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: PCT/US99/28214
; PRIOR FILING DATE: 1999-11-29
; PRIOR APPLICATION NUMBER: PCT/US99/28313
; PRIOR FILING DATE: 1999-11-30
; PRIOR APPLICATION NUMBER: PCT/US99/28564
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/28565
; PRIOR FILING DATE: 1999-12-02
; PRIOR APPLICATION NUMBER: PCT/US99/30095
; PRIOR FILING DATE: 1999-12-16
; PRIOR APPLICATION NUMBER: PCT/US99/30911
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US99/30999
; PRIOR FILING DATE: 1999-12-20
; PRIOR APPLICATION NUMBER: PCT/US00/00219
; PRIOR FILING DATE: 2000-01-05
; NUMBER OF SEQ ID NOS: 423
; SEQ ID NO 212
; LENGTH: 1985
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-905-125A-212

Query Match 8.4%; Score 34; DB 4; Length 1985;
Best Local Similarity 52.1%; Pred. No. 0.95;
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;

QY 2 TGGGAATATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGA 61
Db 1824 TTGGACATAAATGGTGTATGCTTCCGCCAAAAAAGAAAAA 1883

QY 62 AAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTACATTTCAAAAGAACTTC 121
Db 1884 AA 1943

QY 122 AAGATCAAGATAAGAAAGCCCAAGA 147
Db 1944 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 1969

RESULT 14

US-09-902-775A-212
; Sequence 212, Application US/0902775A
; Patent No. 6686451
; GENERAL INFORMATION:
; APPLICANT: Genentech, Inc.
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan L.
; APPLICANT: Ferrara, Napoleone
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gernitsen, Mary E.
; APPLICANT: Goddard, A.
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, Christopher J.
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth, J.

APPLICANT: Kljavin, Ivar J.
APPLICANT: Mather, Jennie P.
APPLICANT: Pan, James
APPLICANT: Paoni, Nicholas F.
APPLICANT: Roy, Margaret Ann
APPLICANT: Stewart, Timothy A.
APPLICANT: Tumas, Daniel
APPLICANT: Williams, P. Mickey
APPLICANT: Wood, William, I.
TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
TITLE OF INVENTION: Acids Encoding the Same
FILE REFERENCE: 10466-14
CURRENT APPLICATION NUMBER: US/09/302,775A
PRIOR APPLICATION NUMBER: PCT/US00/04414
PRIOR FILING DATE: 2001-07-10
PRIOR APPLICATION NUMBER: US 60/143,048
PRIOR FILING DATE: 1999-07-07
PRIOR APPLICATION NUMBER: US 60/145,698
PRIOR FILING DATE: 1999-07-26
PRIOR APPLICATION NUMBER: US 60/146,222
PRIOR FILING DATE: 1999-07-28
PRIOR APPLICATION NUMBER: PCT/US99/20594
PRIOR FILING DATE: 1999-09-08
PRIOR APPLICATION NUMBER: PCT/US99/20944
PRIOR FILING DATE: 1999-09-13
PRIOR APPLICATION NUMBER: PCT/US99/21090
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/21547
PRIOR FILING DATE: 1999-09-15
PRIOR APPLICATION NUMBER: PCT/US99/23089
PRIOR FILING DATE: 1999-10-05
PRIOR APPLICATION NUMBER: PCT/US99/28214
PRIOR FILING DATE: 1999-11-29
PRIOR APPLICATION NUMBER: PCT/US99/28313
PRIOR FILING DATE: 1999-11-30
PRIOR APPLICATION NUMBER: PCT/US99/28564
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/28565
PRIOR FILING DATE: 1999-12-02
PRIOR APPLICATION NUMBER: PCT/US99/30095
PRIOR FILING DATE: 1999-12-15
PRIOR APPLICATION NUMBER: PCT/US99/30911
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US99/30999
PRIOR FILING DATE: 1999-12-20
PRIOR APPLICATION NUMBER: PCT/US00/00219
PRIOR FILING DATE: 2000-01-05
NUMBER OF SEQ ID NOS: 423
SEQ ID NO 212
LENGTH: 1985
TYPE: DNA
ORGANISM: Homo sapiens
US-09-902-775A-212

Query Match 8.4%; Score 34; DB 4; Length 1985;
Best Local Similarity 52.1%; Pred. No. 0.95; Indels 0; Gaps 0;
Matches 76; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 2 TGGGAATTTATCTCTCGGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGA 61
DB 1824 TTGGACAATAAATGGTGCTATGACTGCTTCGGCCAAATAAAAAAAAAAAAAAAAAA 1883
QY 62 AAGGAACCCGATGAGGAAGAAACGGCAGGAATGACTACATTTGAAGAACTTC 121
DB 1884 AA 1943
QY 122 AAGATCAAGATAAGAAAGCCCAAGAA 147
DB 1944 AAAAAAAAAAAAAAAAAAAAAAAAAA 1969

RESULT 15

US-09-627-122-21/c
Sequence 21, Application US/09627122
Patent No. 6472521
GENERAL INFORMATION:
APPLICANT: Uhlmann, Eugen
APPLICANT: Greiner, Beate
APPLICANT: Unger, Eberhard
APPLICANT: Gothe, Gislinde
APPLICANT: Schwerdel, Marc
TITLE OF INVENTION: OLIGONUCLEOTIDES FOR THE INHIBITION OF HUMAN eg5
FILE REFERENCE: 02481.1678
CURRENT APPLICATION NUMBER: US/09/627,122
CURRENT FILING DATE: 2000-07-27
NUMBER OF SEQ ID NOS: 23
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 21
LENGTH: 5340
TYPE: DNA
ORGANISM: Plasmodium falciparum
US-09-627-122-21
Query Match 8.3%; Score 33.8; DB 4; Length 5340;
Best Local Similarity 49.7%; Pred. No. 1.7;
Matches 86; Conservative 0; Mismatches 87; Indels 0; Gaps 0;
QY 1 ATGGGAATTTATCTCTCGGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAG 60
DB 4840 ATGGGNAATTTCTTGTCTCCCATGTTAATTTACACATATATAATATAATAAAT 4781
QY 61 AAGGAACCCGATGAGGAAGAAACGGCAGGAATGACTACATTTGAAGAACTT 120
DB 4780 AAAAAATAATATAAAAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 4721
QY 121 CAAGATCAAGATAAGAAAGCCCAAGAGTTTCATCCACTTCTAATCAGGAAA 173
DB 4720 AAAAAATAATACATATACATATATATATATATATATATATATATATATAT 4668
Search completed: April 22, 2004, 16:09:14
Job time : 86 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 18:31:30 ; Search time 2076 Seconds
(without alignments)
8455.650 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429
Perfect score: 405
Sequence: 1 atgggaattatctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 3364370

Minimum DB seq length: 0
Maximum DB seq length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

GenEmbl:**
1: gb_ba:**
2: gb_htg:**
3: gb_in:**
4: gb_ov:**
5: gb_ov:**
6: gb_pat:**
7: gb_ph:**
8: gb_pl:**
9: gb_pr:**
10: gb_ro:**
11: gb_sts:**
12: gb_sy:**
13: gb_un:**
14: gb_vi:**
15: em_ba:**
16: em_fun:**
17: em_hum:**
18: em_in:**
19: em_mu:**
20: em_om:**
21: em_or:**
22: em_ov:**
23: em_pat:**
24: em_ph:**
25: em_pl:**
26: em_ro:**
27: em_sts:**
28: em_un:**
29: em_vi:**
30: em_htg_hum:**
31: em_htg_inv:**
32: em_htg_other:**
33: em_htg_mus:**
34: em_htg_pln:**
35: em_htg_rod:**
36: em_htg_mam:**
37: em_htg_vrt:**
38: em_sy:**
39: em_htgo_hum:**
40: em_htgo_mus:**
41: em_htgo_other:**

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	43.2	10.7	372	4	S75319
2	41.6	10.3	239	11	G37943
3	39.8	9.8	375	11	BV014350
C 4	38	9.4	255	8	AF351282
C 5	37.2	9.2	283	6	BD211954
6	37	9.1	263	11	CNS06EHA
7	35.6	9.0	404	11	HUMUT5151
8	35.6	8.8	270	8	AY198651
9	34.8	8.6	383	11	HUMUT5116
10	34.8	8.6	399	6	AR417479
11	34.8	8.6	399	6	BD113032
C 12	34.6	8.5	277	8	FSY566324
C 13	34.4	8.5	372	9	BC015361
14	34.4	8.5	397	11	G47173
15	34.2	8.4	353	11	CNS06EJ3
16	34.2	8.4	401	3	AF543047
17	34	8.4	230	11	G50943
C 18	34	8.4	296	10	AY040769S1
19	33.8	8.3	392	3	PFAHMG
20	33.8	8.3	398	17	AF116644
21	33.6	8.3	249	6	AR413685
22	33.6	8.3	249	6	BD109238
23	33.6	8.3	385	3	PFAHMGH5
24	33.4	8.2	240	6	AR413687
25	33.4	8.2	240	6	BD109240
C 26	33.2	8.2	190	9	AY376246
27	33.2	8.2	264	11	HUMUT5168B
28	33.2	8.2	288	11	HUMUT5193B
C 29	33.2	8.2	357	5	ASE276787
30	33	8.1	367	11	HUMUT8147
31	33	8.1	400	6	AR357935
32	32.8	8.1	294	3	AF201952
C 33	32.6	8.0	269	6	AX185675
34	32.4	8.0	310	11	BV077053
35	32.4	8.0	352	3	AY054826
36	32.4	8.0	352	3	AY054852
C 37	32.4	8.0	381	11	BV095162
38	32.2	8.0	120	6	AX353334
39	32.2	8.0	333	6	AX363358
C 40	32	7.9	231	14	D67094
41	32	7.9	273	9	HUMUT8321
42	32	7.9	347	3	PFAHMG
43	31.8	7.9	261	8	AF310859
44	31.8	7.9	302	10	BN000763
C 45	31.8	7.9	384	6	BD244367

ALIGNMENTS

RESULT 1

S75319 372 bp DNA linear NAM 27-MAY-1995
LOCUS 12S rRNA {nuclear clone pNunt.1, tandem repeat macrosatellite}
DEFINITION [Felis catus=domestic cats, lymphocytes, fibroblasts, Genomic, 372 nt].

ACCESSION S75319

VERSION S75319.1 GI:833954

KEYWORDS Felis catus (cat)

SOURCE Felis catus

ORGANISM Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Carnivora; Fissipedia; Felidae; Felis.

REFERENCE 1 (bases 1 to 372)

AUTHORS Lopez,J.V., Yuhki,N., Masuda,R., Modi,W. and O'Brien,S.J.

TITLE Numt, a recent transfer and tandem amplification of mitochondrial DNA to the nuclear genome of the domestic cat
 J. Mol. Evol. 39 (2), 174-190 (1994)
 MEDLINE 95018295
 PUBMED 7932781
 REMARK GenBank staff at the National Library of Medicine created this entry [NCBI gibbon 160474] from the original journal article. This sequence comes from Fig. 6B.
 Map location: D2.

FEATURES

source Location/Qualifiers
 1..372
 /organism="Felis catus"
 /mol_type="genomic DNA"
 /db_xref="taxon:9685"
 gene 1..372
 /gene="12S rRNA"

ORIGIN

Query Match 10.7%; Score 43.2; DB 4; Length 372;
 Best Local Similarity 57.4%; Pred. No. 1.8;
 Matches 78; Conservative 0; Mismatches 58; Indels 0; Gaps 0;
 QY 95 AAATGCTACATTTCAAGAAACCTTCAAGATCAAGATAAGAAAGCCCAAGATTCAT 154
 Db 170 ATATACCGCATCTTCAGCAACCTTAAGAGGAAGAAAGTAGCACAGATCTTAAC 229
 QY 155 CCACCTTTCAATCAGGAACGAGAAATGGCAGTGGTTCTGAAGAGTGTGCTACACTGCA 214
 Db 230 CCAAAAAAGTTAGGTCAAGGTGAGCTCATGAGATGGAGCAATGGCTACACTTCT 289
 QY 215 TTAATCAGATCCCA 230
 Db 290 AAAATAGTAATACCA 305

RESULT 2

G37943 239 bp DNA linear STS 26-FEB-1999
 LOCUS C4M29 Plasmodium falciparum haploid Plasmodium falciparum STS
 DEFINITION genomic, sequence tagged site.
 ACCESSION G37943
 VERSION G37943.2 GI:4337400
 KEYWORDS STS.
 SOURCE Plasmodium falciparum (malaria parasite P. falciparum)
 ORGANISM Plasmodium falciparum
 Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.
 1 (bases 1 to 239)
 Su, X.-z., Ferdig, M.T. and Wellens, T.E.
 Integrated genetic and physical maps of Plasmodium falciparum
 Unpublished (1998)
 TITLE On Mar 6, 1999 this sequence version replaced gi:3056687.
 JOURNAL
 COMMENT

Contact: Thomas E. Wellens

LFD

NIAID

Bldg. 4, Room 126, NIH Campus, Bethesda, MD 20892-0425

Tel: (301) 496-4021

Fax: (301) 402-0079

Email: tew@helix.nih.gov

Primer A: TCTTACACTTTGTCTCATCAT

Primer B: CATTTGGTACTGTAGTGATC

STS size: 239

PCR Profile:

Initial heat: 2min at 94°C

Cycles: 30 cycles, each having steps 20sec at 94°C;

10sec at 45°C; 10sec at 40°C; 30sec at 60°C

Machine: Perkin Elmer GeneAmp 9600

Protocol:

Template: 1-10 ng

Primer: 5 pM each

dNTPs: 0.3 uM of 10 mM stock solution

Taq Polymerase: 0.5 unit

Total Volume: 15 uL

Buffer:

10 mM Tris HCl, pH 8.3
 1.5 mM MgCl2
 50 mM KCl.

FEATURES

source Location/Qualifiers

1..239

/organism="Plasmodium falciparum"

/mol_type="genomic DNA"

/strain="HB3xDd2"

/db_xref="taxon:5833"

/map="4"

/clone_lib="Plasmodium falciparum haploid"

/dev_stage="erythrocytic haploid stage"

1..239

1..20

complement (221..239)

ORIGIN

Query Match 10.3%; Score 41.6; DB 11; Length 239;
 Best Local Similarity 53.8%; Pred. No. 5;
 Matches 86; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
 QY 41 AGAATCAAGAGCCCAAGAAAGGAACCCAGATGAGGAAAGAAACCGCAGGAATCA 100
 Db 56 AAATAAAGAAAGAAATAAAGAAAGAAATAAAGAAAGAAATAAAGAAAGAAATAA 115
 QY 101 CTACATTTGAAAGAAACCTTCAAGATCAAGATAAGAAAGCCCAAGAGTTTCATCCACTT 160
 Db 116 GAAAAATAAAGAAAGAAATAAAGAAAGAAATAAAGAAAGAAATAAAGAAAGAAATAA 175
 QY 161 CTAATCAGGAACGAGAAATGGCAGTGGTTCTCAAGAACT 200
 Db 176 ATAAAAAGAAAGAAAGAAAGGGGGGTAGGGGAAGAACT 215

RESULT 3

BV014350 375 bp DNA linear STS 30-MAY-2003
 LOCUS S208F6381FP2.T0 129S1/SvimJ Mus musculus STS genomic, sequence tagged site.
 DEFINITION BV014350
 ACCESSION BV014350.1 GI:31098245
 VERSION BV014350.1
 KEYWORDS STS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 375)
 Wade, C.M., Kulbokas, E.J. III, Kirby, A.W., Zody, M.C., Mullikin, J.C., Lander, E.S., Lindblad-Toh, K. and Daly, M.J.
 The mosaic structure of variation in the laboratory mouse genome
 Nature 420 (6915), 574-578 (2002)
 TITLE Nature 420 (6915), 574-578 (2002)
 JOURNAL 22354684
 MEDLINE 12466852
 PUBMED

COMMENT

Contact: Kerstin Lindblad-Toh
 Whitehead Institute for Biomedical Research, Center for Genome Research
 320 Charles Street, Cambridge, MA 02141, USA
 Tel: 6172521477
 Fax: 6172580903
 Email: kersli@genome.wi.mit.edu
 Primer A: None
 Primer B: None
 STS size: 375
 Protocol:
 WGS-discovery: Paired-end low-coverage whole genome shotgun reads were generated from 129S1/SvimJ, C3H/HeJ, and BALB/cByJ. The WGS reads were placed uniquely on the MGSCV3 C57BL/6J assembly and SNP detection was carried out by SSAHA-SNP. 225,000 reads were annotated
 as STSs and 81,000 SNPs were annotated with alleles from C57BL/6J

Db 235 AACCTAAACAGCCTTGGGCAGTAGAAGAAGAAGAAAAAGAAGAATAAGAGAAGAAC 176

Qy

KEYWORDS
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
AUTHORS 1 (bases 1 to 270)
TITLE May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R., McCombie, W.R. and Martienssen, R.A.
JOURNAL Arabidopsis genomic sequences flanking Ds enhancer and gene traps in transgenic lines
REFERENCE Unpublished
AUTHORS 2 (bases 1 to 270)
TITLE May, B.P., Smorowski, J., Arroyo, J.-M., Vaughn, M.W., Shen, R., McCombie, W.R. and Martienssen, R.A.
JOURNAL Direct Submission
COMMENT Submitted (17-DEC-2002) Cold Spring Harbor Laboratory, 1 Bungtown Road, Cold Spring Harbor, NY 11724, USA
FEATURES Location/Qualifiers
 source
 1..270
 /organism="Arabidopsis thaliana"
 /mol_type="genomic DNA"
 /db_xref="taxon:3702"
 /clone="ET2091.Ds3.02.07.00.b.270"
 /ecotype="Landsberg"
 /note="transgenic line ET2091"
ORIGIN
 Query Match 8.8%; Score 35.6; DB 8; Length 270;
 Best Local Similarity 51.2%; Pred. No. 1.7e+02;
 Matches 83; Conservative 0; Mismatches 79; Indels 0; Gaps 0;
 QY 19 CGAAATCACTTCCTGGGAGAGATCAAGAGAGCCGAGAGAGAAACCCAGATGAG 78
 Db 58 CAAAAATGAGTTCCTTTAAACAAGATCAATCGTAACCAATCCAAAGCCCAAGAT 117
 QY 79 GAAAGAAACCGCAGGAAATGACTACATTTCAAGAGAAATCTCAAGATCAAGATAGAAA 138
 Db 118 CGAGAGAAACCAAGAGATGATCACTATCTTTAAAGCTTATCGAGTCACTCATCA 177
 QY 139 AGCCAAGAGTTTCATCCATCTTCAATCAGAGAAACGAGAAAT 180
 Db 178 AGCAAAACCCCTACAGAAATTCGAAAGATGAAAGAGGAGAAAT 219
RESULT 9
HUMUT5116 383 bp DNA linear STS 26-JUL-1993
LOCUS Human chromosome 22 STS UT5116, sequence tagged site.
DEFINITION L18521
ACCESSION L18521.1 GI:308187
VERSION STS; PCR primer; STS sequence; microsatellite marker;
KEYWORDS microsatellite repeat; repeat polymorphism; sequence tagged site; tetranucleotide repeat.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
AUTHORS Gaken, S.C., Matsunami, N., Lawrence, E., Carlson, M., Moore, M., Ballard, L., Melis, R., Robertson, M., Bradley, P., Eisner, T., Tangey, A., Rodriguez, P., Albertsen, H., Lalouel, J.-M. and White, R.
TITLE Genetic and physical mapping of simple sequence repeat containing sequence tagged sites from the human genome
JOURNAL Unpublished (1993)
COMMENT Original source text: Homo sapiens DNA.
 Submitted by: Utah Center for Human Genome Research University of Utah, Dept. of Human Genetics
 2160 Eccles Institute of Human Genetics
 Salt Lake City, UT 84112
 e-mail: sts@corona.med.utah.edu
 Primer A: TGCCACTGTACTCCAGCTT

Primer B: AAGATTATTGGCACCAAAACC
 32P-label: A Primer
PCR Profile:
 Initial Denaturation: 94C 300sec
 PCR Cycles: 5
 Denaturation: 94C 10sec
 Annealing: 58C 10sec
 Extension: 72C 20sec
 Mg++: 1mM
 Gel: Acrylamide 7%, Formamide 32%, Urea 34%
Alleles: 4.
FEATURES Location/Qualifiers
 source
 1..383
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 /map="22"
 43..337
 /standard_name="STS UT5116"
 primer_bind 43..61
 primer_bind complement (317..337)
ORIGIN
 Query Match 8.6%; Score 34.8; DB 11; Length 383;
 Best Local Similarity 52.0%; Pred. No. 2.6e+02;
 Matches 78; Conservative 0; Mismatches 72; Indels 0; Gaps 0;
 QY 36 GGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAGAAAGAAACGGCAGGA 95
 Db 122 CGAAGGAG 181
 QY 96 AATGACTACATTTGAAAGAGAAACITCAAGATCAAGATAGAAAGAAAGAAAGAAAGAA 155
 Db 182 AAG 241
 QY 156 CACTTCTTAATCAGGAGAAACGAGAAATGGCAG 185
 Db 242 GCGAGGCGAGTGAGAGAGCGAGGAGGAGAG 271
RESULT 10
AR417479 399 bp DNA linear PAT 18-DEC-2003
LOCUS AR417479
DEFINITION Sequence 8976 from patent US 6639063.
ACCESSION AR417479
VERSION AR417479.1 GI:40172589
KEYWORDS Unknown.
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 399)
AUTHORS Edwards, J.-B.D.M., Jobert, S. and Giordano, J.-Y.
TITLE EST's and encoded human proteins
JOURNAL Patent: US 6639063-A 8976 28-OCT-2003;
FEATURES Location/Qualifiers
 source
 1..399
 /organism="unknown"
 /mol_type="genomic DNA"
ORIGIN
 Query Match 8.6%; Score 34.8; DB 6; Length 399;
 Best Local Similarity 12.4%; Pred. No. 2.6e+02;
 Matches 30; Conservative 110; Mismatches 102; Indels 0; Gaps 0;
 QY 3 GCGAAATATCTCTCGGAAACTCAGTTGCTGCGAGAGAACTCAAAAGAGAGCCCAAGAA 62
 Db 42 KCRERRRRRRWWKSMCMWKKSKSWRSWGMWTKMKGRGAASWAGYNSWWTYTRRWRY 101
 QY 63 AGGAAACCCAGATGAGGAG 122
 Db 102 YFRRKACTKWRAGMWGKAGWAWYAKWYNWNRRTAMKYWAMWMSKRSRRRRAWY 161
 QY 123 AGATCAAGATAAGAAAGCAAGAGAGTTTCATCCACTTCTTAATCAGGAGAAACGAGAG 182

```

162 WNYTMMARRTWGWREASCYRGAYMASAGMYTMMYTMRRKWMYSAGWSMWRKWTRECA 221
183 CAGTGGTTCTCAGGAAGTGCTACACTGTCAATTAATCACATCCCCCATCAGAGATCCTC 242
222 YSCWSSYCWGAKMMYWKTSRWSYSSYRCKYRRSCCWSMSYKWKYYRSWYASC 281
243 CC 244
282 YY 283

RESULT 11
LOCUS BD113032 399 bp DNA linear PAT 18-SEP-2002
DEFINITION EST and encoded human protein.
ACCESSION BD113032
KEYWORDS BD113032.1 GI:320207850
SOURCE JP 2002010789-A/5109
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 399)
AUTHORS Edwards,J.B.D.M., Jobert,S. and Giordano,J.E.
TITLE EST and encoded human protein
JOURNAL Patent: JP 2002010789-A 5109 15-JAN-2002;
GENSET CORP
OS Homo sapiens (human)
PN JP 2002010789-A/5109
PD 15-JAN-2002
PF 07-AUG-2000 JP 2000280989
PR 05-AUG-1999 US 60/147499
PI JEAN BAPTISTE DUMAS MILNE EDWARDS, SEVELIN JOBERT, JEAN EVE PI
GIORDANO
PC C12N15/09, C12N15/09, C07K14/47, C07K16/18, C12N1/15, C12N1/19, PC
C12N1/21,
PC C12N5/10, C12P21/02, C12P21/08, C12Q1/68, C12N15/00, C12N5/00, PC
C12N15/00
CC EST and encoded human protein
FH Key
FT source
FT source 1..399
Location/Qualifiers
/organism="Homo sapiens (human)"
/ol_type="genomic DNA"
/db_xref="taxon:9606"

FEATURES
source
1..399
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"

ORIGIN
Query Match 8.6%; Score 34.8; DB 6; Length 399;
Best Local Similarity 12.4%; Pred. No. 2.6e+02;
Matches 30; Conservative 110; Mismatches 102; Indels 0; Gaps 0;

QY 3 GGGAAATATCTCCGCGAAACTCAGTTGCTGGAGAGAAATCAAAAGAGAGCCCAAGAA 62
DB 42 KRRRRRAWWKSKWKKSKSWRSGWMTKRWKGRGAASWAGYMSWMTYTRWRY 101
QY 63 AGGAACCCAGATGAGGAAAGAAACCGCAGGAATGACTATGTAAGAAAGAACTCA 122
DB 102 YRYRKAATKWRAGWGWKAGWAWAYAKWYTWMTWRTAMKYWAMKSKSRMRRRAWYA 161
QY 123 AGATCAAGATAGAAAGCCAGAGATTCTATCCACTTAATCAGGAACAGAGATGG 182
DB 162 WNYTMMARRTWGWREASCYRGAYMASAGMYTMMYTMRRKWMYSAGWSMWRKWTRECA 221
QY 183 CAGTGGTTCTCAGGAAGTGCTACACTGTCAATTAATCACATCCCCCATCAGAGATCCTC 242
DB 222 YSCWSSYCWGAKMMYWKTSRWSYSSYRCKYRRSCCWSMSYKWKYYRSWYASC 281
243 CC 244
282 YY 283
```

```

RESULT 12
LOCUS FSV566324/c 277 bp DNA linear PLN 30-JUN-2003
DEFINITION Ficus sycamorus microsatellite DNA, clone FIGT09D12.
ACCESSION AJ566324
VERSION AJ566324.1 GI:32399995
KEYWORDS microsatellite; repetitive DNA.
SOURCE Ficus sycamorus
ORGANISM Ficus sycamorus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; Core eudicots;
Rosids; eurosids 1; Rosales; Moraceae; Ficus.
REFERENCE
1
AUTHORS Dawson,D.A., Ahmed,S., Compton,S.G., Gilmartin,P.M. and Burke,T.
TITLE Isolation of microsatellite loci in the fig, Ficus sycamorus
JOURNAL Unpublished
KEYWORDS 2 (bases 1 to 277)
AUTHORS Dawson,D.A.
TITLE Direct Submission
JOURNAL Submitted (28-MAY-2003) Dawson D.A., Department of Animal and Plant
Sciences, University of Sheffield, Alfred Denny Building, Western
Bank, Sheffield, S10 2TN, UNITED KINGDOM
FEATURES
source
1..277
Location/Qualifiers
/organism="Ficus sycamorus"
/mol_type="genomic DNA"
/db_xref="taxon:182129"
/clone="FIGT09D12"
87..190
satellite
1..277
note="microsatellite
rpt_unit:(T)6(C)1(T)6(G)1(T)7(A)2(TTTC)1(TTTC)1
(TTTC)5(TTTC)7(C)1(T)18"
/rpt_type="TANDEM"

ORIGIN
Query Match 8.5%; Score 34.6; DB 8; Length 277;
Best Local Similarity 58.1%; Pred. No. 3.1e+02;
Matches 61; Conservative 0; Mismatches 44; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAGAGGAAACCCAGATGAGGAAGAAACGCGAGAAAT 98
DB 188 AAAAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAA 129
QY 99 GACTACATTGAAAGAAACTTCAAGATCAAGATCAAGATCAAGAGAAAGCCCA 143
DB 128 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAAA 84

RESULT 13
LOCUS BC015361 372 bp mRNA
DEFINITION Homo sapiens, clone IMAGE:3886095, mRNA.
ACCESSION BC015361
VERSION BC015361.1 GI:15929880
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens (human)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE
1 (bases 1 to 372)
AUTHORS Strausberg,R.
TITLE Direct Submission
JOURNAL Submitted (01-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT Contact: MGC help desk
Email: gcgabs-r@mail.nih.gov
Tissue Procurement: DCID/DPF/Gazdar
CDNA Library Preparation: Life Technologies, Inc.
```

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Sequencing Group at the Stanford Human Genome
 Center, Stanford University School of Medicine, Stanford, CA 94305
 Web site: <http://www-shgc.stanford.edu>
 Contact: (Dickson, Mark) md@paxil.stanford.edu
 Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
 R. M.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAX Plate: 22 Row: c Column: 19.

FEATURES

Location/Qualifiers
 1..372
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:3886095"
 /tissue_type="Lung, carcinoma, large cell
 undifferentiated"
 /clone_lib="NIH_MGC_69"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6"

ORIGIN

Query Match 8.5%; Score 34.4; DB 9; Length 372;
 Best Local Similarity 52.9%; Pred. No. 3.3e+02;
 Matches 74; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 8 ATTATCTCTCGAAACTAGTTCCTGGAGAGATCAAGAACCCCAAGAGGAA 67
 Db 232 ACTTTGACCTATTAAAAAGGTGTACCACTAAAAAAGAAAAAAGAAAAA 291
 QY 68 ACCAGATGAGGAAGAAACGGCAGGAATGACTACATTGGAAGAAACTTCAAGATC 127
 Db 292 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 351
 QY 128 AAGATAGAAAGCAAGCAAGAA 147
 Db 352 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 371

RESULT 14

G47173
 LOCUS 397 bp DNA linear STS 23-MAR-1999
 DEFINITION Z15801.1 Zebrafish AB Danio rerio STS genomic clone Z15801 5',
 sequence tagged site.
 ACCESSION G47173
 VERSION G47173.1 GI:4493512
 KEYWORDS STS.
 SOURCE Danio rerio (zebrafish)
 ORGANISM Danio rerio
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
 Cypriniformes; Cyprinidae; Danio.
 1 (bases 1 to 397)
 Ziniti, J., Jackson, D.G. and Fishman, M.C.
 A zebrafish microsatellite map
 Unpublished (1999)

REFERENCE

AUTHORS Contact: Mark C. Fishman
 TITLE Cardiovascular Research Center
 JOURNAL Massachusetts General Hospital
 COMMENT Mail code 1494100A, 149 13th Street, Charlestown, MA 02129, USA
 Fax: 617/265806
 Email: fishman@gh.cvr.harvard.edu
 http://zebrafish.mgh.harvard.edu
 Primer A: CCGCAAGTACCCAAATAA
 Primer B: AGCGTTTCTCCAGTCAA
 STS size: 223
 PCR Profile:
 Presoak: 94 degrees C for 5.0 minutes
 Denaturation: 94 degrees C for 1.0 minute
 Annealing: 58 degrees C for 1.0 minute

Polymerization: 72 degrees C for 1.5 minute
 PCR Cycles: 27
 Thermal Cycler: MJ Research PTC-100

Protocol:

Template: 10 ng
 Primer: each 375 nM
 dNTPs: each 200 uM
 Taq Polymerase: 0.034 units/ul
 Total Vol: 10 ul

Buffer:

MgCL2: 1.5 mM
 KCl: 50 mM
 Tris-HCl: 10 mM
 pH: 8.3.

FEATURES

Location/Qualifiers
 1..397
 /organism="Danio rerio"
 /mol_type="genomic DNA"
 /strain="AB"
 /db_xref="taxon:7955"
 /clone="Z15801"
 /sex="F"
 /clone_lib="Zebrafish AB"
 /dev_stage="Adult"
 /lab_host="DH5alpha/IO"
 /note="Vector: M13mp19 with added BstXI site; V-type:
 phage; Genomic DNA from a single adult Zebrafish of AB
 strain was digested with AluI, Cac81, HaeIII, NlaVI, or
 RsaI. Fragments in the range of 250-500 bp were gel
 purified and a BstXI linker was added. The fragments were
 cloned into a modified M13mp19 vector and transformed
 into E. Coli DH5alpha. Microsatellite sequences were
 screened with labeled d(CA)15 and d(GT)15 oligonucleotide
 probes."

STS
 primer_bind 86..308
 primer_bind 86..105
 complement(289..308)

ORIGIN

Query Match 8.5%; Score 34.4; DB 11; Length 397;
 Best Local Similarity 52.9%; Pred. No. 3.3e+02;
 Matches 74; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
 QY 206 ACACCTGTCTATTATCACATCCCATCAGATCTCCTGAGTCCCAATGATGGCT 265
 Db 51 ACACCTATTCAATGACATCTTATGACGACATCACCGCAAGTACCCAAAATAATCTCT 110
 QY 266 ATGAGACATTCAGTCCCTCACAGGAAGTGAGACAGTTTAGAAGAGGTGAGACAG 325
 Db 111 GTGTGTGTTTGTAGTCTTCTAAATGAAGTAGGAGATTCAAACAATCTGGGTGAGG 170
 QY 326 AATATGCCCTTCTTAGGACT 345
 Db 171 AATTAACCATTTGTCGTCT 190

RESULT 15

CNS06EJ3
 LOCUS 353 bp DNA linear STS 10-JAN-2001
 DEFINITION T7 end of clone AF0AA022A02 of library AF0AA from strain CBS 732 of
 Zygosaaccharomyces rouxii, sequence tagged site.
 ACCESSION AL395173
 VERSION AL395173.1 GI:12146610
 KEYWORDS STS.
 SOURCE Zygosaaccharomyces rouxii
 ORGANISM Zygosaaccharomyces rouxii
 Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 Saccharomycetales; Saccharomycetaceae; Zygosaaccharomyces.
 1 (bases 1 to 353)
 Souciet, J.L., Aigle, M., Artiguenave, F., Blandin, G.,
 Bolotin-Fukuhara, W., Bon, E., Brottier, P., Casaregola, S.,
 de-Montigny, J., Dujon, B., Durrens, P., Lepingle, A., Llorente, B.,

Malpertuy,A., Neuvesglise,C., Ozier-Kalogeropoulos,O., Potier,S.,
 Saurin,W., Tekala,F., Toffano-Nioche,C., Wesolowski-Louvel,M.,
 Wincker,P. and Weissenbach,J.
 Genomic exploration of the hemiascomycetous yeasts: 1. A set of
 yeast species for molecular evolution studies
 FEBS Lett. 487 (1), 3-12 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20584711
 11152876

REFERENCE
 2 (bases 1 to 353)
 de Montigny,J., Straub,M., Potier,S., Tekala,F., Dujon,B.,
 Wincker,P., Artiguenave,F. and Souciet,J.
 Genomic exploration of the hemiascomycetous yeasts: 8.
 Zygosccharomyces rouxii
 FEBS Lett. 487 (1), 52-55 (2000)

JOURNAL
 MEDLINE
 PUBMED
 20584718
 11152883

REFERENCE
 3 (bases 1 to 353)
 Genoscope.
 Direct Submission
 Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,
 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)

JOURNAL
 MEDLINE
 PUBMED
 20584718
 11152883

REFERENCE
 3 (bases 1 to 353)
 Genoscope.
 Direct Submission
 Submitted (06-SEP-2000) Genoscope - Centre National de Sequencage,
 2 rue Gaston Cremieux, CP 5706, 91057 EVRY cedex, FRANCE. (E-mail :
 seqrefgenoscope.cns.fr - Web : www.genoscope.cns.fr)

COMMENT
 This STS is part of a random genomic sequencing program of thirteen
 yeast species: Saccharomyces bayanus var. uvarum, Saccharomyces
 exiguus, Saccharomyces servazzii, Zygosccharomyces rouxii,
 Saccharomyces kluyveri, Kluyveromyces thermotolerans, Kluyveromyces
 lactis var. lactis, Kluyveromyces marxianus var. marxianus, Pichia
 angusta, Debaryomyces hansenii var. hansenii, Pichia sorbitophila,
 Candida tropicalis and Yarrowia lipolytica. Genomic inserts of 3 to
 5 Kb were prepared and both extremities were sequenced. See
 keywords for description of this sequence and for the sequence of
 the other extremity of this insert.

FEATURES
 source
 1..353
 /organism="Zygosccharomyces rouxii"
 /mol_type="genomic DNA"
 /strain="CBS 732"
 /db_xref="taxon:4956"
 /clone="AR0AA022A02"
 /clone_lib="AR0AA"
 /note="end : T7"
 1..353

STS
 ORIGIN

Query Match 8.4%; Score 34.2; DB 11; Length 353;
 Best Local Similarity 54.3%; Pred. No. 3.8e+02;
 Matches 69; Conservative 0; Mismatches 58; Indels 0; Gaps 0;

QY 21 AAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGAGCCCAAGAAAGGAAACCCAGATGAGGA 80
 |||||
 Db 153 AAAACAAAACATTAATAAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 212
 |||||

QY 81 AAGAAACGGCAGGAATGACTACATTTTGAAGAGAAACTTCAAGATCAAGATAGAAAG 140
 |||||
 Db 213 AAAAAAATTCAAAAAAGAAAAAATAGGAAAAAAGAAAAAAGAAATTAATAAAAAAAGAAAAA 272
 |||||

QY 141 CCAAGAA 147
 |||||

Db 273 AAAAAA 279

Search completed: April 23, 2004, 03:08:42
 Job time : 2081 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 18:29:17 ; Search time 2022 Seconds

(without alignments)

5981.297 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429

Perfect score: 405

Sequence: 1 atgggaaattatctctgcg.....atgaagtgtgtttccacac 405

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 16189338

Minimum DB seq length: 0

Maximum DB seq length: 405

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST*

1: em_estba:**

2: em_estum:**

3: em_estin:**

4: em_estmd:**

5: em_estov:**

6: em_estpl:**

7: em_estro:**

8: em_estc:**

9: gb_estl:**

10: gb_est2:**

11: gb_estc:**

12: gb_est3:**

13: gb_est4:**

14: gb_est5:**

15: em_estfun:**

16: em_estom:**

17: em_gss_hum:**

18: em_gss_inv:**

19: em_gss_pln:**

20: em_gss_vrt:**

21: em_gss_fun:**

22: em_gss_mam:**

23: em_gss_mus:**

24: em_gss_pro:**

25: em_gss_rod:**

26: em_gss_phg:**

27: em_gss_vrl:**

28: gb_gss1:**

29: gb_gss2:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54.8	16.0	288	10 AW480906	AW480906 33516 MAR
2	46.8	11.6	353	29 AY404587	AY404587 Pan trogl
3	46.6	11.5	259	29 CE192340	CE192340 tigr-gss-
4	44	10.9	203	29 CE091328	CE091328 tigr-gss-

5	43.4	10.7	404	13	BX367419
6	42	10.4	345	14	CF351413
7	41.8	10.3	160	28	AZ273050
8	41.8	10.3	274	29	CE329306
9	41.2	10.2	357	28	AZ419893
10	40.6	10.0	276	28	AQ645726
11	40.6	10.0	359	13	EX460850
12	40.4	10.0	244	9	AL725744
13	40.2	9.9	340	29	CNS020HP
14	39.8	9.8	261	29	CNS0232W
15	39.6	9.8	372	9	AU039357
16	39.4	9.7	261	29	CNS006BK
17	39.4	9.7	364	29	CE343646
18	39.4	9.7	395	28	AZ937971
19	39.4	9.7	399	14	CD677582
20	39.2	9.7	356	14	CF379377
21	39	9.6	358	14	CF803201
22	39	9.6	401	9	AL134217
23	38.8	9.6	210	9	AA225339
24	38.8	9.6	364	29	CE462660
25	38.6	9.5	252	12	EG981526
26	38.6	9.5	268	12	BI493055
27	38.6	9.5	382	12	EG497771
28	38.6	9.5	389	10	AW786169
29	38.6	9.5	400	12	EG382689
30	38.4	9.5	244	29	CE825890
31	38.4	9.5	279	12	BI671444
32	38.4	9.5	292	14	CD422569
33	38.4	9.5	293	12	EM155090
34	38.4	9.5	320	13	EX361849
35	38.4	9.5	355	29	CE797846
36	38.2	9.4	186	13	C93279
37	38.2	9.4	313	29	CNS00CZT
38	38.2	9.4	327	29	CE476461
39	38.2	9.4	357	29	CE302878
40	38	9.4	131	14	CF330236
41	38	9.4	210	12	BM873785
42	38	9.4	250	14	CE289843
43	38	9.4	278	14	CF544822
44	38	9.4	304	10	BB291389
45	38	9.4	383	14	CF804314

ALIGNMENTS

RESULT 1
AW480906 288 bp mRNA linear EST 09-JUL-2000
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
COMMENT

AW480906 288 bp mRNA linear EST 09-JUL-2000
33516 MARC 2PIG Sus scrofa cDNA 5', mRNA sequence.
AW480906.1 GI:7050949
EST.
Sus scrofa (pig)
Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.
1 (bases 1 to 288)
Fahrenkrug,S.C., Smith,T.P.L., Freking,B.A., Cho,J., White,J.,
Vallet,J., Wise,T., Rohrer,G.A., Pertea,G., Sultana,R.,
Quackenbush,J. and Keele,J.W.
Porcine gene discovery by normalized cDNA-library sequencing and
EST cluster assembly
Mamm. Genome 13 (8), 475-478 (2002)
22213789
12226715
Contact: Smith TPL
USDA, ARS, US Meat Animal Research Center
PO Box 166, Clay Center, NE 68933-0166, USA
Tel: 402 762 4366
Fax: 402 762 4390
Email: smith@email.marc.usda.gov
Single pass sequencing. Bases called and trimmed with phred

v0.980904.e. Vector identified by cross_match with the -minscore 20 and -minmatch 12 options.

PCR Primers
 FORWARD: AGGAACAGCTATGACCAT
 BACKWARD: GTTTCCTCAGTCACGACG
 Plate: 19 row: F column: 21
 Seq primer: ATTAGTGACACATATAG.
 Location/Qualifiers
 1..288
 /organism="Sus scrofa"
 /mol_type="mRNA"
 /db_xref="taxon:9823"
 /tissue_type="pooled"
 /lab_host="DHIOB"
 /clone_lib="MARC 2P1G"
 /note="vector: pcMV SPORT6; Site 1: NotI; Site 2: SalI;
 Library made from pooled tissue from testis, ovary,
 endometrium, hypothalamus, pituitary, and placenta."

FEATURES

source

ORIGIN

Query Match 16.0%; Score 64.8; DB 10; Length 288;
 Best Local Similarity 75.8%; Pred. No. 4e-05;
 Matches 94; Conservative 0; Mismatches 27; Indels 3; Gaps 1;
 QY 285 CACAAGGAAGTACAGAGTTTACAGAAAGTACAGACAGATATGCCCTTCTTAGGAC 344
 DB 2 CACAAGAGAGTGAGACCAATTAAGAGGAGGATCAGAACAGATATGCCCTTCTTAGGAC 61
 QY 345 TTCGTAGT---AGGCTTCTTCTGACCCATGAGCATGATTATGAAGTTGTGTTCC 401
 DB 62 GACTTACACTCCAAAGTTTTTCTCTATACCCCTGAGATGATTGAAGTTGTGTTCC 121
 QY 402 ACAC 405
 DB 122 TCAC 125

RESULT 2

AY404587

LOCUS

DEFINITION Pan troglodytes GCET2 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.

ACCESSION

AY404587

VERSION

AY404587.1

GI:39760564

KEYWORDS

GSS.

SOURCE

Pan troglodytes (chimpanzee)

ORGANISM

REFERENCE

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

Inferring nonneutral evolution from human-chimp-mouse orthologous

TITLE

Gene trios

JOURNAL

Science 302 (5652), 1960-1963 (2003)

PUBMED

14671302

REFERENCE

AUTHORS

Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejariwal, A.,
 Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B.,
 Ferriera, S., Wang, G., Zheng, X.H., White, T.J., Sninsky, J.J.,
 Adams, M.D. and Cargill, M.

Direct Submission

TITLE

JOURNAL

Submitted (16-NOV-2003)

Rockville, MD 20850, USA

COMMENT

This sequence was made by sequencing genomic exons and ordering

them based on alignment.

FEATURES

source

1..353

/organism="Pan troglodytes"

/mol_type="genomic DNA"

/db_xref="taxon:9598"

gene

ORIGIN

<1..>353

/gene="GCET2"

/locus_tag="HCM1940"

Query Match

Best Local Similarity

Matches 106; Conservative

0; Mismatches

113; Indels

3; Gaps

1;

QY

114

AAAACCTCAAGATCAAGATAAGAAAGCCAAAGAGTTTTCATCCACTTCTATCAGAAAAA 173

DB

131

AAAGANNNAAGATTCCCAAAACGAAATGAAGGAATGTCACTCACTCCATCCAGNNNN 190

QY

174

CGAAGATGGCAGTGGTCTCTGAAGAAGTGTCTACACTGTCTATTAATCAACA---TCCCCCA 230

DB

191

NNNNNNNNNNNNNNNNNNAGNAGCTGTCTANANCNCTCATNNNCGGGTCTCTG 250

QY

231

TCAGAGATCCTCCCTGAGCTCCAATGATGATGGGTATGAGAACATTGACTCCCTCACAAG 290

DB

251

NACRAGGCCATCAGGGAACCTCTGCTGAAGACTACTATGAGATGTTCCTCGAAGCTGA 310

QY

291

GAAAGTGAAGATTTAGAGAAAGTCAAGACAGATATGTC 332

DB

311

GAGACCCAGAGAGCCCTTGGGAGGAACCTGAGACTGAGTATTC 352

RESULT 3

CE192340/c

LOCUS

tigr-gss-dog-1700037146986

Dog Library

Canis familiaris genomic,

genomic survey sequence.

CE192340

VERSION

CE192340.1

GI:35347993

KEYWORDS

GSS

SOURCE

Canis familiaris (dog)

ORGANISM

REFERENCE

AUTHORS

Kirknes, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
 Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
 Venter, J.C.

The dog genome: survey sequencing and comparative analysis

TITLE

JOURNAL

Science 301 (5641), 1898-1903 (2003)

MEDLINE

22875432

PUBMED

14512627

COMMENT

Contact: Kirkness EF

The Institute for Genomic Research

Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,

Rockville, MD 20850, USA

Tel: 301-838-0200

Fax: 301-838-0208

Email: ekirknes@tigr.org

Class: shotgun.

FEATURES

Location/Qualifiers

1..259

/organism="Canis familiaris"

/mol_type="genomic DNA"

/strains="Standard Poodle"

/db_xref="taxon:9615"

/clone_lib="Dog Library"

/note="Site 1: BstXI; Libraries were prepared from

peripheral blood"

ORIGIN

Query Match

Best Local Similarity

Matches 79; Conservative

0; Mismatches

54; Indels

0; Gaps

0;

QY

47

AAAGAAGCCCAAGAAAGAAACCCAGATGAGAAAGAAACCGGAGGAATGACTACAT 106

DB

166

AAAGAAGCCCAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 107

```
QY 107 TTGAAGAAGAACTTCAAGATCAAGATAAGAGAAAGCCAGAGAGTTTCATCCATCTTCTTAATC 166
Db 106 AAGAAAGAAAGAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAGAGAAAG 47
QY 167 AGGAAACAGAGAA 179
Db 46 AAGAAAGAGAGAA 34

RESULT 4
CE091328 203 bp DNA linear GSS 24-SEP-2003
LOCUS tigr-gss-dog-17000359492028 Dog Library canis familiaris genomic,
DEFINITION genomic survey sequence.
ACCESSION CE091328
VERSION CE091328.1 GI:35158174
KEYWORDS GSS.
SOURCE Canis familiaris (dog)
ORGANISM Canis familiaris
REFERENCE 1 (bases 1 to 203)
AUTHORS Kirness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
Venter,J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
FEATURES
source
1..203
Location/Qualifiers
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="Standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"
/note="Site 1: BstXI; Libraries were prepared from
peripheral blood"
ORIGIN
Query Match 10.9%; Score 44; DB 29; Length 203;
Best Local Similarity 55.1%; Pred. No. 3.4;
Matches 86; Conservative 0; Mismatches 70; Indels 0; Gaps 0;
QY 20 GAAAGAACTCAGTTGCTGGGAGAGATCAAGAGAGCCAGAGAGAAAGCCAGATGAGG 79
Db 20 GAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 79
QY 80 AAGAGAAACGGCAGAGAAATGACTACATTTGAAAGAAAGAACTTCAAGATCAAGATAAGAAA 139
Db 80 AAGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 139
QY 140 GCCAGAGAGTTTCATCCACTTCTAATCAGGAGAAACG 175
Db 140 GAAAGAGAAATGTTTATTCCTCTGGAGAGAGAAATG 175

RESULT 5
BX367419 404 bp mRNA linear EST 08-MAY-2003
LOCUS BX367419
DEFINITION BX367419 Homo sapiens T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED
ACCESSION BX367419
VERSION BX367419.1 GI:30455633
KEYWORDS EST.
```

```
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Mammalia; Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 404)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished (2001)
COMMENT Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: genref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 10480.f For
more information about this cluster, see
http://www.genoscope.cns.fr/
cgi-bin/cluster.cgi?seq=CS0AP004ZG06NP1&cluster=10480.f. Contact :
Peng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0AP004ZG06NP1.
Location/Qualifiers
1..404
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ009YM22"
/cell_type="T CELLS (JURKAT CELL LINE) COT 10-NORMALIZED"
/cell_line="JURKAT"
/clone_lib="Homo sapiens T CELLS (JURKAT CELL LINE) COT
10-NORMALIZED"
/note="1st strand cDNA was primed with a NotI-oligo(dT)
primer. Five prime end enriched, double-strand cDNA was
digested with Not I and cloned into the Not I and EcoR V
sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN
Query Match 10.7%; Score 43.4; DB 13; Length 404;
Best Local Similarity 55.0%; Pred. No. 5;
Matches 71; Conservative 6; Mismatches 52; Indels 0; Gaps 0;
QY 19 CGAAAGAACTCAGTTGCTGGGAGAGATCAAGAGAGCCAGAGAGAAAGCCAGATGAG 78
Db 191 CCCGACTCAGATGCTGAGAAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 250
QY 79 GAAAGAAACGGCAGAGAAATGACTACATTTGAAAGAAAGAACTTCAAGATCAAGATAAGAAA 138
Db 251 AAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 310
QY 139 AGCCAGAGAA 147
Db 311 AAGAGAGAGAA 319

RESULT 6
CF351413 345 bp mRNA linear EST 20-AUG-2003
LOCUS lab89el2.y1 SIEP Mus musculus CDNA 5', mRNA sequence.
DEFINITION CF351413
ACCESSION CF351413.1 GI:33954424
VERSION CF351413.1
KEYWORDS EST.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 345)
AUTHORS Tidwell,R., Clifton,S., Marra,M., Hillier,L., Pape,D., Martin,J.,
Wylie,T., Theising,B., Bowers,Y., Gibbons,M., Ritter,E., Bennet,J.,
Ronko,I., Tsagarishvili,R., Belaygorod,L., Grow,A., Maguire,L.,
Waterston,R. and Wilson,R.
TITLE WashU Stem cell EST Project
JOURNAL Unpublished (2002)
COMMENT Contact: Jeff Gordon and Mike Lovett
WashU, Human Genetics Division
```

Washington University School of Medicine
Library materials provided by: Stappenbeck Th.S. and Gordon J.I.
Library constructed by: Korshunova I. and Lovett M. DNA sequencing
by: Washington University Genome Sequencing Center For information
on obtaining a clone please contact: Rose Tidwell
(rtidwell@genetics.wustl.edu)
Putative full length read
vector to vector length is
Seq primer: -40RP from Gibco.
Location/Qualifiers
1. .345
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
/tissue_type="small intestinal epithelial progenitors"
/lab_host="DMSalpha Ultra Max cells (Invitrogen)"
/clone_lib="slep"
/notes="vector: pAMP1; The library was synthesized with
modified SMART primers with dUTP at the end. After
treatment with UDG the cDNA was cloned in pAMP1 vector by
annealing."

FEATURES

source

ORIGIN

Query Match 10.4%; Score 42; DB 14; Length 345;
Best Local Similarity 52.9%; Pred. No. 11;
Matches 90; Conservative 0; Mismatches 80; Indels 0; Gaps 0;
QY 14 TCCTGCGAAACTCAGTTCCTGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAG 73
Db 5 TGCTATTACCTCTACTAAGGGGGAAGAAAAAAGAAAAAAGAAAAAAGAAAA 64
QY 74 ATGAGGAAGAAACGGCAGGAATGACTATTTGAAGAAACCTTCAGATCAAGATA 133
Db 65 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 124
QY 134 AGAAAGCCAGAGATTCTATCCACTTCTATCAGGAAGAGCAATGGC 183
Db 125 AAAAAAAGAAAAAGTCCCTCCCTTTTACCCCAAGAAAAAGCC 174

RESULT 7
LOCUS AZ273050 160 bp DNA linear GSS 26-JUL-2000
DEFINITION RPCI-23-85A5 TJC RPCI-23 Mus musculus genomic clone RPCI-23-85A5,
genomic survey sequence.
ACCESSION AZ273050
VERSION AZ273050.1 GI:9486590
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 160)
AUTHORS Zhao, S., Nierman, W., Feldblyum, T., Malek, J., Shatsman, S.,
Akinret, B., Levins, M., McGann, S., Teegaye, G., Geer, K., Krol, M., de
Jong, P. and Fraser, C.M.
TITLE Mouse BAC End Sequences from Library RPCI-23
JOURNAL Mouse BAC End Sequences from Library RPCI-23
COMMENT Other GSSs: RPCI-23-85A5.TV
Contact: Shaying Zhao
Department of Eukaryotic Genomics
The Institute for Genomic Research
9712 Medical Center Dr., Rockville, MD 20850, USA
Tel: 301 838 0200
Fax: 301 838 0208
Email: szhao@tigr.org
Clones are derived from the mouse BAC library RPCI-23. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (<http://bacpac.med.buffalo.edu/orderingframe.htm>)
or from Resea ch Genetics (info@resgen.com). BAC end page:
http://www.tigr.org/tldb/bacends/mouse/bac_end_intro.html
Plate: 85 row: A column: 5

Seq primer: sp6

Class: BAC ends.

FEATURES

source

Location/Qualifiers
1. .160
/organism="Mus musculus"
/mol_type="genomic DNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clones="RPCI-23-85A5"
/sex="Female"
/lab_host="DH10B"
/clone_lib="RPCI-23"
/note="Organ: Kidney/Brain; Vector: pBACe3.6; Site_1:
ECORI; Site_2: EcoRI; Female C57BL/6J mouse kidney and/or
brain genomic DNA was isolated and partially digested
with a combination of EcoRI and EcoRI Methyase. Size
selected DNA was cloned into the pBACe3.6 vector at the
ECORI sites. The ligation products were transformed into
DH10B electrocompetent cells (BRL Life Technologies)."

ORIGIN

Query Match 10.3%; Score 41.8; DB 28; Length 160;
Best Local Similarity 58.4%; Pred. No. 11;
Matches 73; Conservative 0; Mismatches 52; Indels 0; Gaps 0;
QY 22 AAATCTAGTTCCTGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATCAGGAA 81
Db 36 AAATCTGGGGGTGGGGGGAGTCAAAAGAGAGAGAGAGAGAGAGAGAGAGAGAG 95
QY 82 AGAAAGCCGAGGAATGACTACTATTTGAAAGAGAACTTCAAGATCAAGATAAGAAAGC 141
Db 96 AAGAAAG 155
QY 142 CAAGA 146
Db 156 AAGA 160

RESULT 8

CE329306

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

COMMENT

CE329306 274 bp DNA linear GSS 26-SEP-2003
tigr-gss-dog-17000333970436 Dog Library Canis familiaris genomic,
genomic survey sequence.
CE329306
CE329306.1 GI:36144095
GSS.
Canis familiaris (dog)
Canis familiaris
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Carnivora; Fissipedia; Canidae; Canis.
REFERENCE 1 (bases 1 to 274)
AUTHORS Kirkness, E.F., Bafna, V., Halpern, A.L., Levy, S., Remington, K.,
Rusch, D.B., Delcher, A.L., Pop, M., Wang, W., Fraser, C.M. and
Venter, J.C.
TITLE The dog genome: survey sequencing and comparative analysis
JOURNAL Science 301 (5641), 1898-1903 (2003)
MEDLINE 22875432
PUBMED 14512627
COMMENT Contact: Kirkness EF
The Institute for Genomic Research
Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
Rockville, MD 20850, USA
Tel: 301-838-0200
Fax: 301-838-0208
Email: ekirknes@tigr.org
Class: shotgun.
Location/Qualifiers
1. .274
/organism="Canis familiaris"
/mol_type="genomic DNA"
/strain="standard Poodle"
/db_xref="taxon:9615"
/clone_lib="Dog Library"

FEATURES

source

ALL78961.1 GI:7817018
 GDS; genome survey sequence.
 Tetraodon nigroviridis
 Tetraodon nigroviridis
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
 Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
 Tetraodontoidea; Tetraodontidae; Tetraodon.

1 Roest Crolius,H., Jaillon,O., Dasilva,C., Bouneau,L., Fisher,C.,
 Bernot,A., Fizames,C., Wincker,P., Brottier,P., Quetier,F.,
 Saurin,W. and Weissenbach,J.
 Estimate of human gene number provided by genome-wide analysis
 using Tetraodon nigroviridis DNA sequence
 Nat. Genet. 25 (2), 235-238 (2000)
 20296633
 10835645

2 Roest Crolius,H., Jaillon,O., Dasilva,C., Ozouf-Costaz,C.,
 Fizames,C., Fischer,C., Bouneau,L., Billault,A., Quetier,F.,
 Saurin,W., Bernot,A. and Weissenbach,J.
 Characterization and repeat analysis of the compact genome of the
 freshwater pufferfish Tetraodon nigroviridis
 Genome Res. 10 (7), 939-949 (2000)
 20359837
 10899143

3 (bases 1 to 261)
 Genoscope.
 Direct Submission
 Submitted (12-APR-2000) Genoscope - Centre National de Sequences :
 BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
 - Web : www.genoscope.cns.fr)
 This sequence is a single read and was generated as part of a large
 scale clone-end sequencing project of the Tetraodon nigroviridis
 genome. For more information, please take a look at
 http://www.genoscope.cns.fr/Tetraodon.

FEATURES
 source
 1..261
 /location=Qualifiers
 /organism="Tetraodon nigroviridis"
 /mol_type="genomic DNA"
 /db_xref="taxon:99883"
 /clone="232A01"
 /clone_lib="G"
 /note="Genoscope sequence ID : C0AG232AA01SPI-end :
 PUC-ori"

ORIGIN
 Query Match 9.8%; Score 39.8; DB 29; Length 261;
 Best Local Similarity 51.88; Pred. No. 35;
 Matches 73; Conservative 4; Mismatches 64; Indels 0; Gaps 0

QY 39 AGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGAAAGAAACGCGCAGGAAT 98
 12 AAAGAAAGAAATGEBNNWGA 71
 QY 99 GACTACATTTCAAGAGAAACTTCAAGATCAAGATAGAAAGCCAGAGAGTTTCATCCAC 158
 Db 72 AAATAAATAGAAAAAATTTAAAAAATAAAAAAAAAAAAAAAAAAATGNTAWAAAAA 131
 QY 159 TTCTAATCAGAAAAACGAGAA 179
 Db 132 ATTTAANCAGNAAAGAAA 152

RESULT 15
 AU039357
 LOCUS
 DEFINITION
 AU039357 Dictyostelium discoideum SL (H.Urushihara) Dictyostelium
 discoideum cDNA clone SLH629, mRNA sequence.
 ACCESSION
 AU039357.1 GI:4008597
 VERSION
 EST.
 KEYWORDS
 Dictyostelium discoideum

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 17:20:54 ; Search time 291 Seconds
(without alignments)
5912.441 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429
Perfect score: 405
Sequence: 1 atgggaattatctctgog.....atgaagttgtgttccacac 405

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 4428564

Minimum DB seq length: 0
Maximum DB seq length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N Geneseq_29Jan04:*

- 1: geneseqn1980s.*
- 2: geneseqn1990s.*
- 3: geneseqn2000s.*
- 4: geneseqn2001as.*
- 5: geneseqn2001bs.*
- 6: geneseqn2002s.*
- 7: geneseqn2003as.*
- 8: geneseqn2003bs.*
- 9: geneseqn2003cs.*
- 10: geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	38.8	9.6	172	5 AAF24365	Aaf24365 Retrovira
2	38.6	9.5	403	7 ABT22783	Abt22783 Breast ca
3	37.2	9.2	283	2 AAZ12627	Aaz12627 Human gen
4	37.2	9.2	384	5 ABV18980	Abv18980 Human pro
5	36.6	9.0	390	4 AAI86998	Aai86998 Human pol
6	36.4	9.0	325	5 AAS29049	Aas29049 CDNA enco
7	36.4	9.0	325	6 ABS68189	Abs68189 CDNA enco
8	36.4	9.0	325	9 ADC25183	Adc25183 Human cdn
9	36.2	8.9	346	5 ABV54289	Abv54289 Human pro
10	36.2	8.9	396	4 AAI85197	Aai85197 Human pol
11	36	8.9	343	7 ABX39335	Abx39335 Bovine ES
12	36	8.9	381	6 ABL78974	Ab178974 Human ova
13	36	8.9	382	5 ABV54300	Abv54300 Human pro
14	35.8	8.8	317	5 ABV58181	Abv58181 Human pro
15	35.6	8.8	330	5 ABV18573	Abv18573 Human pro
16	35.4	8.7	329	5 ABV50079	Abv50079 Human pro
17	35.4	8.7	332	7 ABX46789	Abx46789 Bovine ES
18	35.4	8.7	341	7 ABX47137	Abx47137 Bovine ES
19	35.2	8.7	385	4 AAI80160	Aai80160 Human pol
20	35	8.6	209	5 ABV57456	Abv57456 Human pro
21	35	8.6	233	5 ABV58773	Abv58773 Human pro
22	34.8	8.6	338	5 ABV48535	Abv48535 Human pro
23	34.8	8.6	385	5 ABV54241	Abv54241 Human pro

C	24	34.8	8.6	402	4	AAI82045	Aai82045 Human pol
	25	34.4	8.5	195	7	ACA25512	Aca25512 Prokaryot
	26	34.2	8.4	252	5	ABV58215	Abv58215 Human pro
	27	34.2	8.4	290	6	ABL87783	Ab187783 Human ova
	28	34.2	8.4	299	7	ABX38127	Abx38127 Bovine ES
	29	34.2	8.4	309	7	ABX41877	Abx41877 Bovine ES
C	30	34.2	8.4	393	4	AAI11189	Aai11189 Human bre
	31	34.2	8.4	394	4	AAI87163	Aai87163 Human pol
	32	34.2	8.4	403	4	AAI87975	Aai87975 Human pol
	33	34.2	8.4	403	4	AAI83197	Aai83197 Human pol
	34	34	8.4	318	8	ADA58930	Ada58930 Maize suc
	35	34	8.4	366	4	AAI83212	Aai83212 Human pol
	36	34	8.4	392	7	ABX39643	Abx39643 Bovine ES
	37	33.8	8.3	208	3	ACQ98557	Acq98557 Human col
	38	33.8	8.3	224	5	ABV57906	Abv57906 Human pro
	39	33.8	8.3	262	5	ABV58178	Abv58178 Human pro
C	40	33.8	8.3	332	7	ABX39719	Abx39719 Bovine ES
	41	33.8	8.3	347	5	ABV58528	Abv58528 Human pro
	42	33.8	8.3	355	4	AAI81491	Aai81491 Human pol
C	43	33.8	8.3	357	4	AAK90758	Aak90758 Human dig
	44	33.8	8.3	357	4	AAK90757	Aak90757 Human dig
C	45	33.8	8.3	357	4	AAK90760	Aak90760 Human dig

ALIGNMENTS

RESULT 1

AAF24365
ID AAF24365 standard; DNA; 172 BP.
XX
AC AAF24365;
XX
DT 24-APR-2001 (first entry)
XX
DE Retroviral recombination assay coding sequence fragment #10.
XX
KW Retroviral recombination assay; gene therapy vector; viral vector; gag;
KW pol; replication; HIV; ds.
XX
OS Unidentified.
XX
PN WO200104360-A2.
XX
PD 18-JAN-2001.
XX
PF 06-JUL-2000; 2000WO-US018597.
XX
PR 09-JUL-1999; 99US-0143015P.
PR 10-NOV-1999; 99US-0164626P.
XX
XX (UABR-) UAB RES FOUND.
XX (TRAN-) TRANZYME INC.
PI Kappes JC, Wu X, Wakefield J;
XX
XX WFI; 2001-091927/10.
DR
XX
XX Retroviral recombination assays, systems and cells, useful for evaluating the risk of producing a replication-competent retrovirus from a retroviral-based vector.
XX
XX Example 5; Fig 14; 88pp; English.
XX
XX The present invention describes a method for detecting a retroviral genetic recombinant having gag and pol functions, involving a cell suspected of having a recombinant and propagating the recombinant in the presence of any necessary helper functions. This is useful as it enables the detection of vectors which are able to replicate in the host cell, and allows the production of vectors suitable for gene therapy. The present sequence is a DNA fragment used to demonstrate the method of the invention
XX

CC cancerous state of a mammalian cell, comprising detecting at least one
CC differentially expressed gene product in a test sample from a cell
CC suspected of being cancerous, where the gene product is encoded by one of
CC the 5248 polynucleotide sequences given in AA212532 to AA217779. The
CC polynucleotides can be used as a source of primers and probes, which can
CC be used for a variety of purpose, e.g. detection of expression levels,
CC mapping, tissue typing or profiling, forensic, genetic analysis and
CC detection of polymorphisms. Polypeptides encoded by the polynucleotides
CC can be used for raising antibodies for experimental, diagnostic and
CC therapeutic purposes. The polynucleotides may also be used to construct
CC arrays for diagnostics (which may be used to determine function of an
CC encoded protein); and to detect differences in expression levels between
CC two cells (e.g. to identify abnormal or diseased tissue in a human, to
CC identify a genetic predisposition or susceptibility to a disease such as
CC cancer). The polynucleotides of the invention are especially used in the
CC diagnosis, prognosis and management of colorectal cancer, breast cancer,
CC and lung cancer. The polynucleotides can also be used to screen for
CC peptide analogues and antagonists
XX
SQ Sequence 283 BP; 20 A; 27 C; 34 G; 126 T; 0 U; 76 Other;

Query Match 9.2%; Score 37.2; DB 2; Length 283;
Best Local Similarity 42.9%; Pred. No. 1;
Matches 66; Conservative 0; Mismatches 88; Indels 0; Gaps 0;
QY 41 AGAATCAAGAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACGGCAGGAAATGA 100
Db 255 AAANGNAGANNCAN 196
QY 101 CTACATTTGAAGAAACTTCAGATCAAGATAGAAAGCCAGAGATTCATCCACTT 160
Db 195 AAATAAANGNAAAT 136
QY 161 CTAATCAGGAAAGCAGAGATGGCAGTGGTCTTGA 194
Db 135 AAANNCAGNTANGTGA 102

RESULT 4
ABV18980
ID ABV18980 standard; cDNA; 384 BP.
XX
AC ABV18980;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 18971.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
FN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
EF 20-FEB-2001; 2001WO-US005171.
XX
FR 17-FEB-2000; 2000US-0183319P.
FR 16-MAR-2000; 2000US-0189862P.
FR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
XX Novel isolated nucleic acid molecule associated with cancerous state of

PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
PS Claim 1; Page 3116; 11750pp; English.
XX
CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (I) is also useful as a pharmacodynamic or pharmacogenomic marker
XX
SQ Sequence 384 BP; 163 A; 57 C; 23 G; 82 T; 0 U; 59 Other;

Query Match 9.2%; Score 37.2; DB 5; Length 384;
Best Local Similarity 50.6%; Pred. No. 1.1;
Matches 78; Conservative 0; Mismatches 76; Indels 0; Gaps 0;
QY 8 ATTATCTCTCGGAAACTCAGTTCCTCGGAGAGATCAAAAGAGCCCAAGAAAGAA 67
Db 1 AGTACTCTTAGAATTAATTAAGCGGGGGGAGACAAAANANANANANANANANANAN 60
QY 68 ACCCAGATGAGGAAAGAAACCGCAGGAAATGACTACATTTGAAAGAAAACTTCAAGATC 127
Db 61 AAANA 120
QY 128 AAGATAAGAAAGCCCAAGAGTTTCATCCACTTC 161
Db 121 AAANA 154

RESULT 5
AA186998
ID AA186998 standard; cDNA; 390 BP.
XX
AC AA186998;
XX
DT 06-NOV-2001 (first entry)
XX
DE Human polynucleotide SEQ ID NO 7058.
XX
KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
OS Homo sapiens.
XX
FN WO200164835-A2.
XX
PD 07-SEP-2001.
XX
EF 26-FEB-2001; 2001WO-US004927.
XX
FR 28-FEB-2000; 2000US-00515126.
PR 18-MAY-2000; 2000US-00577409.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Tang YT, Liu C, Drmanac RT;
XX
DR WPI; 2001-514838/56.
DR P-PSDB; AAO07067.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
PT and treating e.g. leukemia, inflammation and immune disorders.
XX

PS Claim 1; SEQ ID NO 7058; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and

CC the encoded proteins (AAO0010-AAO1910) that exhibit activity relating to

CC cytokine, cell proliferation or cell differentiation or which may induce

CC production of other cytokines in other cell populations. The

CC polynucleotides and polypeptides are useful in gene therapy, vaccines or

CC peptide therapy. The polypeptides have various cytokine-like activities,

CC e.g. stem cell growth factor activity, haematopoiesis regulating

CC activity, tissue growth factor activity, immunomodulatory activity and

CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and

CC inflammation. Note: The sequence data for this patent did not form part

CC of the printed specification, but was obtained in electronic format

CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX Sequence 390 BP; 229 A; 48 C; 49 G; 57 T; 0 U; 7 Other;

SEQ

Query Match 9.0%; Score 36.6; DB 4; Length 390;

Best Local Similarity 54.0%; Pred. No. 1.7;

Matches 75; Conservative 0; Mismatches 64; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAGAAAGCCAGATGAGGAAAGAAACGGCAGGAAAT 98

Dd 135 AA 194

QY 99 GACTACATTGAAAGAAACTTCAAGATCAAGATAAGAAAGCCAGAGTTTCATCCAC 158

Dd 195 AA 254

QY 159 TTCTAATCAGGAACGAG 177

Dd 255 AATAAATAAAAAAAAAAGGGG 273

RESULT 6

ID AAS29049

XX AAS29049 standard; cDNA; 325 BP.

AC AAS29049;

XX

DT 21-NOV-2001 (first entry)

XX

DE cDNA encoding for human DNA-binding protein #20.

DE

DE Human; DNA-binding protein; histone; chromo domain protein;

KW Chromatin organisation modifier; Y-box binding protein; DNA organisation;

KW gene transcription; malignant disease; autoimmune disorder;

KW rheumatic disease; genetic abnormality; infectious disease;

KW neurological disorder; gene therapy; immunomodulatory; anti-HIV;

KW anti rheumatic; anti microbial; cytostatic; ss.

XX

OS Homo sapiens.

XX

PN WO20015162-A1.

XX

PD 02-AUG-2001.

XX

PF 17-JAN-2001; 2001WO-US001305.

XX

XX 31-JAN-2000; 2000US-0179065P.

PR 04-FEB-2000; 2000US-0180628P.

PR 24-FEB-2000; 2000US-0184664P.

PR 02-MAR-2000; 2000US-0186350P.

PR 16-MAR-2000; 2000US-0189874P.

PR 17-MAR-2000; 2000US-0190076P.

PR 18-APR-2000; 2000US-0198123P.

PR 19-MAY-2000; 2000US-0205515P.

PR 07-JUN-2000; 2000US-0209467P.

PR 28-JUN-2000; 2000US-0214886P.

PR 30-JUN-2000; 2000US-0215135P.

PR 07-JUL-2000; 2000US-0215647P.

PR 07-JUL-2000; 2000US-0216880P.

PR

PR 11-JUL-2000; 2000US-0217487P.

PR 11-JUL-2000; 2000US-0217496P.

PR 14-JUL-2000; 2000US-0218290P.

PR 26-JUL-2000; 2000US-0220963P.

PR 26-JUL-2000; 2000US-0220964P.

PR 14-AUG-2000; 2000US-0224518P.

PR 14-AUG-2000; 2000US-0224519P.

PR 14-AUG-2000; 2000US-0225213P.

PR 14-AUG-2000; 2000US-0225214P.

PR 14-AUG-2000; 2000US-0225266P.

PR 14-AUG-2000; 2000US-0225267P.

PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.

PR 14-AUG-2000; 2000US-0225447P.

PR 14-AUG-2000; 2000US-0225757P.

PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.

PR 18-AUG-2000; 2000US-0226279P.

PR 22-AUG-2000; 2000US-0226681P.

PR 22-AUG-2000; 2000US-0226689P.

PR 22-AUG-2000; 2000US-0227182P.

PR 23-AUG-2000; 2000US-0227009P.

PR 30-AUG-2000; 2000US-0228924P.

PR 01-SEP-2000; 2000US-0229287P.

PR 01-SEP-2000; 2000US-0229343P.

PR 01-SEP-2000; 2000US-0229344P.

PR 01-SEP-2000; 2000US-0229345P.

PR 05-SEP-2000; 2000US-0229509P.

PR 05-SEP-2000; 2000US-0229513P.

PR 06-SEP-2000; 2000US-0230437P.

PR 06-SEP-2000; 2000US-0230438P.

PR 08-SEP-2000; 2000US-0231242P.

PR 08-SEP-2000; 2000US-0231243P.

PR 08-SEP-2000; 2000US-0231244P.

PR 08-SEP-2000; 2000US-0231413P.

PR 08-SEP-2000; 2000US-0231414P.

PR 08-SEP-2000; 2000US-0232080P.

PR 08-SEP-2000; 2000US-0232081P.

PR 12-SEP-2000; 2000US-0231968P.

PR 14-SEP-2000; 2000US-0232397P.

PR 14-SEP-2000; 2000US-0232398P.

PR 14-SEP-2000; 2000US-0232399P.

PR 14-SEP-2000; 2000US-0232400P.

PR 14-SEP-2000; 2000US-0232401P.

PR 14-SEP-2000; 2000US-0232403P.

PR 14-SEP-2000; 2000US-0233064P.

PR 14-SEP-2000; 2000US-0233065P.

PR 21-SEP-2000; 2000US-0234223P.

PR 21-SEP-2000; 2000US-0234274P.

PR 25-SEP-2000; 2000US-0234997P.

PR 25-SEP-2000; 2000US-0234998P.

PR 26-SEP-2000; 2000US-0235484P.

PR 27-SEP-2000; 2000US-0235834P.

PR 27-SEP-2000; 2000US-0235836P.

PR 29-SEP-2000; 2000US-0236327P.

PR 29-SEP-2000; 2000US-0236367P.

PR 29-SEP-2000; 2000US-0236368P.

PR 29-SEP-2000; 2000US-0236369P.

PR 29-SEP-2000; 2000US-0236370P.

PR 02-OCT-2000; 2000US-0236802P.

PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.

PR 02-OCT-2000; 2000US-0237039P.

PR 02-OCT-2000; 2000US-0237040P.

PR 13-OCT-2000; 2000US-0239935P.

PR 13-OCT-2000; 2000US-0239937P.

PR 20-OCT-2000; 2000US-0240960P.

PR 20-OCT-2000; 2000US-0241221P.

PR 20-OCT-2000; 2000US-0241785P.

PR 20-OCT-2000; 2000US-0241786P.

PR 20-OCT-2000; 2000US-0241787P.

PR 20-OCT-2000; 2000US-0241808P.

PR 20-OCT-2000; 2000US-0241809P.

20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0246617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 01-DEC-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-02559678P.
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX PI Rosen CA, Barash SC, Ruben SM;
XX
XX WPI; 2001-465557/50.
XX P-PSDB; AAU18173.
XX
XX Nucleic acid molecules encoding human secreted chromosomal binding
XX proteins, used in preventing, treating or ameliorating a disorder, e.g.
XX Alzheimer's and Parkinson's diseases and cancers.
XX
XX Claim 4; SEQ ID NO 30; 561pp; English.
XX
XX The present invention relates to the isolation of novel DNA-binding
XX proteins (AAU18154-AAU18281), and cDNA and genomic sequences encoding for
XX these proteins. DNA-binding proteins such as histones, chromatin
XX organisation modifier domain proteins, and Y-box binding proteins may
XX contribute to diseases resulting from aberrant DNA organisation and/or
XX gene transcription. The sequences of the invention are useful in
XX screening assays to identify antagonists and/or agonists that may enhance
XX or block activities mediated by DNA-binding proteins. Blockers of DNA-
XX binding proteins may be useful in treating disorders such as malignant

diseases (e.g. cancer), autoimmune disorders (e.g. diabetes mellitus),
rheumatic diseases (e.g. rheumatoid arthritis), genetic abnormalities
(e.g. cystic fibrosis), infectious diseases (e.g. HIV) and neurological
disorders (e.g. Alzheimer's disease). The polynucleotide sequences of the
invention may also be used in gene therapy. AAS29030-AAS29157 represent
cDNA sequences encoding for novel DNA-binding proteins. Note: The
sequence data for this patent did not form part of the printed
specification, but was obtained in electronic format directly from WIPO
at ftp.wipo.int/pub/published_pct_sequences

Sequence 325 BP; 174 A; 38 C; 36 G; 77 T; 0 U; 0 Other;
Query Match 9.0%; Score 36.4; DB 5; Length 325;
Best Local Similarity 53.5%; Pred. No. 1.8;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 6 AAATTATCTCTCGGAAACTCAGTTGCTGGAGAGAAATCAAAAGAGCCCAAGAGG 65
DB 182 AGTTATCTCTTTCGAAAGAGATATTTCAACCTAAATAAAAAAAAAAAAAAAAAA 241
QY 66 AAACCCAGATGAGGAAAGAAACGGCAGGAAATGACTCATTTGAAAGAACTTCAAGA 125
DB 242 AA 301
QY 126 TCAAGATTAAGAAAGCCAGAA 147
DB 302 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 323

RESULT 7
ABS68189
ID ABS68189 standard; cDNA; 325 BP.
XX
XX AC ABS68189;
XX
XX DT 18-NOV-2002 (first entry)
XX
XX DE cDNA encoding human DNA-binding protein #20.
XX
XX KW Human; DNA-binding protein; B cell immunodeficiency; autoimmune disorder;
XX severe combined immunodeficiency; rheumatoid arthritis; Crohn's disease;
XX diabetes mellitus; allergy; asthma; inflammatory condition; thrombosis;
XX graft-versus-host disease; blood-related disorder; atherosclerosis;
XX hyperproliferative disorder; cancer; renal disorder; arrhythmia;
XX acute glomerulonephritis; cardiovascular disorder; respiratory disorder;
XX Goodpasture's syndrome; neurological disorder; Addison's disease;
XX Parkinson's disease; endocrine disorder; infectious disease;
XX reproductive system disorder; endometriosis; viral infection; vaccine;
XX viral infection; bacterial infection; fungal infection; gene therapy; ss.
XX gastrointestinal disorder; multiple sclerosis; gene therapy; ss.
XX
XX OS Homo sapiens.
XX
XX PN US2002102638-A1.
XX
XX XX 01-AUG-2002.
XX
XX PF 17-JAN-2001; 2001US-00764846.
XX
XX XX 31-JAN-2000; 2000US-0179065P.
XX PR 04-FEB-2000; 2000US-0180628P.
XX PR 28-JUN-2000; 2000US-0214886P.
XX PR 07-JUL-2000; 2000US-0216847P.
XX PR 07-JUL-2000; 2000US-0216880P.
XX PR 11-JUL-2000; 2000US-0217487P.
XX PR 11-JUL-2000; 2000US-0217496P.
XX PR 14-JUL-2000; 2000US-0218290P.
XX PR 26-JUL-2000; 2000US-0220963P.
XX PR 26-JUL-2000; 2000US-0220964P.
XX PR 14-AUG-2000; 2000US-0224518P.
XX PR 14-AUG-2000; 2000US-0224519P.
XX PR 14-AUG-2000; 2000US-0225267P.
XX PR 14-AUG-2000; 2000US-0225268P.

PR 14-AUG-2000; 2000US-0225270P.
PR 14-AUG-2000; 2000US-0225447P.
PR 14-AUG-2000; 2000US-0225757P.
PR 14-AUG-2000; 2000US-0225758P.
PR 22-AUG-2000; 2000US-0226868P.
PR 30-AUG-2000; 2000US-0226924P.
PR 01-SEP-2000; 2000US-0228287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 08-SEP-2000; 2000US-0231413P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 27-SEP-2000; 2000US-0235834P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239335P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX (ROSE/) ROSEN C A.
PA (RUBE/) RUBEN S M.
PA (BARA/) BARASH S C.
XX
XX Rosen CA, Ruben SM, Barash SC;
XX
XX WPI: 2002-690611/74.
DR P-PSDB; ABG92594.
XX

Novel DNA-binding protein useful for diagnosis, prognosis, prevention and treatment of immune, hyperproliferative, respiratory, cardiovascular, reproductive, endocrine, gastrointestinal and neurological disorders.

Claim 1; SEQ ID NO 30; 225pp; English.

The present invention relates to a new DNA-binding protein. The invention is useful in treating, preventing, diagnosing and/or prognosing immunodeficiencies (e.g. B cell immunodeficiencies, severe combined immunodeficiencies), autoimmune disorders (rheumatoid arthritis, multiple sclerosis, diabetes mellitus), allergic reactions and conditions (e.g. asthma), inflammatory conditions, graft-versus-host disease, blood-related disorders (thrombosis, atherosclerosis), hyperproliferative disorders (e.g. cancer), renal disorders (e.g. acute glomerulonephritis), cardiovascular disorders (e.g. arrhythmia), respiratory disorders (Goodpasture's syndrome), neurological disorders (e.g. Alzheimer's disease, Parkinson's disease), endocrine disorders (e.g. Addison's disease), reproductive system disorders (e.g. endometriosis), infectious diseases (e.g. viral, bacterial or fungal infections) and gastrointestinal disorders (e.g. Crohn's disease). The invention is also useful to stimulate neuronal growth and treat, prevent, and/or diagnose neuronal damage which occurs in certain neuronal disorders or neurodegenerative conditions. The present nucleic acid sequence encodes a human DNA-binding protein of the invention. Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from USPTO at <http://seqdata.uspto.gov/sequence>

XX SQ Sequence 325 BP; 174 A; 38 C; 36 G; 77 T; 0 U; 0 Other;
Query Match 9.0%; Score 36.4; DB 6; Length 325;
Best Local Similarity 53.5%; Pred. No. 1.8;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;
QY 6 AAATTATCTCTCGGAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGG 65
DB 182 AGATTATCTTATTCGAAAAGATATTTCAAACTAAATAAAAAAAAAAAAAAAA 241
QY 66 AAACCCAGATGAGGAGAAACGCGAGGAATGACTACTTGAAGAGAACTTCAAGA 125
DB 242 AA 301
QY 126 TCAGATAAGAAAGCCAGAA 147
DB 302 AAAAAAAAAAAAAAAAAAAAAAAAAAAAA 323
RESULT 8
ADC25183
ID ADC25183 standard; cDNA; 325 BP.
XX
AC ADC25183;
XX
DT 18-DEC-2003 (first entry)
XX
DE Human cDNA from extracellular matrix gene 20.
XX
KW Extracellular matrix protein; cytostatic; antibacterial; virucide;
KW neuroprotective; gynaecological; gastrointestinal-Gen; cardiant;
KW cardiovascular-Gen; nephrotropic; antiinflammatory; muscular-Gen;
KW respiratory-Gen; immunosuppressive; cerebroprotective; vasotropic;
KW nootropic; antiallergic; cancer; bacterial infection; viral infection;
KW neural disorder; immune system disorder; blood disorder;
KW muscular disorder; reproductive disorder; gastrointestinal disorder;
KW pulmonary disorder; cardiovascular disorder; renal disorder;
KW inflammatory disorder; proliferative disorder; human; gene therapy; ss;
KW gene.
XX
XX Homo sapiens.
XX
XX US2003049650-A1.
XX
XX 13-MAR-2003.
XX
XX 07-MAR-2002; 2002US-00091483.
XX
PR 31-JAN-2000; 2000US-0179065P.
PR 04-FEB-2000; 2000US-0180628P.
PR 24-FEB-2000; 2000US-0184664P.
PR 02-MAR-2000; 2000US-0186350P.
PR 16-MAR-2000; 2000US-0189874P.
PR 17-MAR-2000; 2000US-0190076P.
PR 18-APR-2000; 2000US-0198123P.
PR 19-MAY-2000; 2000US-0205515P.
PR 07-JUN-2000; 2000US-0209467P.
PR 28-JUN-2000; 2000US-0214886P.
PR 30-JUN-2000; 2000US-0215135P.
PR 07-JUL-2000; 2000US-0216647P.
PR 07-JUL-2000; 2000US-0216880P.
PR 11-JUL-2000; 2000US-0217487P.
PR 11-JUL-2000; 2000US-0217496P.
PR 14-JUL-2000; 2000US-0218290P.
PR 26-JUL-2000; 2000US-0220963P.
PR 26-JUL-2000; 2000US-0220964P.
PR 14-AUG-2000; 2000US-0224518P.
PR 14-AUG-2000; 2000US-0224519P.
PR 14-AUG-2000; 2000US-0225213P.
PR 14-AUG-2000; 2000US-0225214P.
PR 14-AUG-2000; 2000US-0225266P.
PR 14-AUG-2000; 2000US-0225267P.

CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,

	Query Match	9.0%;	Score 36.4;	DB 9;	Length 325;
	Best Local Similarity	53.5%;	Pred. No. 1.8;		
	Matches 76;	Conservative 0;	Mismatches 66;	Indels 0;	Gaps 0;
QY	6	AATATTATCTCGTCGCAAAACTCAGTTGCCTGGGAGAGAAATCAAAGAAGCCCAAGGAAGG	65		
Db	182	AGATTATCTATTGCAAAAAGATATTTCAACCTTAATAAAAAAAAAAAAAAAAAAAAAA	241		
QY	66	AAACCAGCATGAGGAAGAAACCGCAGGAATGACTTACATTTTGAAGGAAAACTTCAAGA	125		
Db	242	AA	301		
QY	126	TCAAGATAGAAAGCCCAAGAA	147		
Db	302	AAAAAAAAAAAAAAAAAAAAAAAAAAAA	323		

RESULT 9
ABV54289
ID ABV54289 standard: cDNA: 346 BP.

XX	Sequence	346 BP;	138 A;	85 C;	52 G;	71 T;	0 U;	0 Other;	
SQ	Query Match	8.9%;	Score	36.2;	DB	5;	Length	346;	
	Best Local Similarity	53.1%;	Prod. No.	2.1;	68;	Indels	0;	Gaps	0
	Matches	77;	Conservative	0;	Mismatches				
QY	14	TCCTGCGAAA	ACTCAGT	TGCTGCTGGGAGAGAA	TCAA	AGAAAGCC	CAAGAACCC	CAG	73
DB	14	TACTTCTAGAA	TATTA	TTAAACGCGGGG	AAAA	AAAA	AAAA	AAAA	73
QY	74	ATGAGGAAG	AAAAACGCGCGG	AAATGACTAC	ATTTC	CAAGAAAA	ACTTCAAGAT	CAAGATA	133
DB	74	AGTAAAA	AAAA	AAAA	AAAA	AAAA	AAAA	AAAA	133
QY	134	AGAAAGCC	AAAGATTTCAT	TCCAC					158
DB	134	AAAA	AAAA	AAAA	AAAA	AGGGG	CCCCCCC		158

RESULT 10
AAI85197
ID AAI85197 standard; cDNA; 396 BP.
XX
AC AAI85197;
XX
XX 06-NOV-2001 (first entry)
DT
DT
DE Human polynucleotide SEQ ID NO 5257.
XX
XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
KW tissue growth factor; immunomodulatory; cancer; leukaemia;
KW nervous system disorders; arthritis; inflammation; ss.
XX
XX Homo sapiens.
OS
XX WO200164835-A2.
FN
XX
XX 07-SEP-2001.
PD
XX
XX 26-FEB-2001; 2001WO-US004927.
PP
XX
XX 28-FEB-2000; 2000US-00515126.
PR
XX
XX 18-MAY-2000; 2000US-00577409.
PR
XX
XX (HYSE-) HYSEQ INC.
PA
XX
XX Tang YT, Liu C, Drmanac RT;
PI
XX
XX WPI: 2001-514838/56.
DR
DR P-PSDB; AA005366.
XX
XX Isolated nucleic acids and polypeptides, useful for preventing diagnosis
PT and treating e.g. leukemia, inflammation and immune disorders.
PT
XX
XX Claim 1; SEQ ID NO 5257; 1399pp + Sequence Listing; English.
PS
XX
XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
CC the encoded proteins (AAO00010-AAO13910) that exhibit activity elating to
CC cytokine, cell proliferation or cell differentiation or which may induce
CC production of other cytokines in other cell populations. The
CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
CC peptide therapy. The polypeptides have various cytokine-like activities,
CC e.g. stem cell growth factor activity, haematopoiesis regulating
CC activity, tissue growth factor activity, immunomodulatory activity and
CC activin/inhibin activity and may be useful in the diagnosis and/or
CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 396 BP; 178 A; 47 C; 100 G; 71 T; 0 U; 0 Other;

CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,

Query Match 9.0%; Score 36.4; DB 9; Length 325;
Best Local Similarity 53.5%; Pred. No. 1.8;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 6 AATATTCTCTCGGAAACTCAGTTGCCTGGGAGAGATCAAAAGAAGCCCAAGGAAGG 65
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 182 AGATTATCTATTGCAAAAAGATATTTCACCTTAAAAAATAAAAAAAAAAAAAA 241
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 66 AAACCCRGATGAGGAAGAAACCGCAGGAATGACTACATTTGAAAGAAAACTTCAAGA 125
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 242 AA 301
| | | | | | | | | | | | | | | | | | | | | | | | | | |
QY 126 TCAAGTAGAAAAAGCCAGAA 147
| | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 302 AAAAAAAAAAAAAAAAAAAAAAAAAA 323
| | | | | | | | | | | | | | | | | | | | | | | | | | |

RESULT 9
ABV54289
ABV54289 standard; cDNA; 346 BP.
AC ABV54289;
XX XX
DT DT
XX XX
DE DE
XX XX
KW KW
XX XX
OS OS
FN FN
XX XX
PD PD
XX XX
PF PF
XX XX
PR PR
PR PR
PR PR
PR PR
PR PR
PR PR
PR PR
PA PA
XX XX
PI PI
XX XX
DR DR
XX XX
PT PT
PT PT
PT PT
XX XX
PS PS
XX XX

17-SEP-2002 (first entry)
Human prostate expression marker cDNA 54280.
Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
pharmacogenomic marker; gene; ss.
Homo sapiens.
WC200160860-A2.
23-AUG-2001.
20-FEB-2001; 2001WO-US005171.
17-FEB-2000; 2000US-0183313P.
16-MAR-2000; 2000US-0189862P.
25-MAY-2000; 2000US-0207454P.
09-JUN-2000; 2000US-0211314P.
18-JUL-2000; 2000US-0213007P.
13-DEC-2000; 2000US-0255281P.

(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
Schlegel R, Endege WO, Monahan JE;
WPI; 2001-662795/76.

Novel isolated nucleic acid molecule associated with cancerous state of prostate cells and correlating with presence of prostate cancer, useful for detecting presence of prostate cancer, stage of prostate cancer.

Claim 1; Page 10498; 11750pp; English.

The invention relates to an isolated nucleic acid molecule (I) comprising a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the specification or its complement. (II) is useful for: (a) assessing whether a patient is afflicted with prostate cancer; (b) monitoring the progression of prostate cancer in a patient; (c) assessing the efficacy of a test compound to inhibit prostate cancer in a patient; (d) assessing the efficacy of a therapy for inhibiting prostate cancer in a patient; (e) selecting a composition for inhibiting prostate cancer in a patient; (f) assessing the prostate cell carcinogenic potential of a compound; (g) determining whether prostate cancer has metastasized in a patient; (h) assessing the aggressiveness or indolence of prostate cancer in a patient; (i) is also useful as a pharmacodynamic or pharmacogenomic marker

CC assay and identifying the protein in the supernatant having the
CC activity). The nucleic acids and proteins display the following
CC activities Cytostatic, antibacterial, Virucide, Neuroprotective,

RESULT 9
ABV54289
ID ABV54289 standard: cDNA: 346 BP.

Query Match 8.9%; Score 36.2; DB 4; Length 396;
 Best Local Similarity 56.2%; Pred. No. 2.2;
 Matches 68; Conservative 0; Mismatches 53; Indels 0; Gaps 0;

QY 20 GAAACTCAGTTGCTGGGAGAGATCAAAAGAGCCCAAGAGGAAAGCCAGATGAGG 79
 DB 35 GCAACCTGATGAAGCTACAAAAAACAACCAAGGAAAAAAGAAAAAAGAA 94
 QY 80 AAGCAAAACGCGAGGAATGACTACATTGCAAGAGAACTTCAAGATCAAGATAAGAAAA 139
 DB 95 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 154

QY 140 G 140
 DB 155 G 155

RESULT 11
 ABX39335
 ID ABX39335 standard; cDNA; 343 BP.
 XX AC ABX39335;
 XX DT 20-FEB-2003 (first entry)
 XX DE Bovine EST associated with lactation/muscle/fat deposition #4500.
 XX KW Bovine; ss; EST; expressed sequence tag; lactation; LMFD;
 XX KW muscle deposition; fat deposition; genome mapping; gene identification;
 XX KW gene analysis; cattle breeding.
 XX OS Bos Taurus.
 XX XN US2002137139-A1.
 XX PD 26-SEP-2002.
 XX PF 24-SEP-2001; 2001US-00960352.
 XX PR 12-JAN-1999; 99US-0115707P.
 XX PR 11-JAN-2000; 2000US-00480902.
 XX PA (BYAT/) BYATT J C.
 XX PA (MATH/) MATHIALAGAN N.
 XX PA (TAON/) TAO N.
 XX PA (WARR/) WARREN W C.
 XX PI Byatt JC, Mathialagan N, Tao N, Warren WC;
 XX WPI; 2003-110599/10.
 XX PT New nucleic acid associated with lactation, and muscle and fat
 XX PT deposition, useful for genome mapping, gene identification and analysis,
 XX PT cattle breeding, or for genetically improving cattle.
 XX PS Claim 2; SEQ ID NO 4500; 245pp; English.

The invention relates to a purified nucleic acid molecule associated with
 lactation or muscle and fat deposition (designated LMFD), derived from
 cattle, and the LMFD nucleic acid can specifically hybridize to a second
 nucleic acid molecule comprising any of 15112 nucleotide sequences,
 appearing as ABX34836-ABX49947, or complements of them. Also included are
 ; (1) a transformed cell having a nucleic acid comprising an LMFD nucleic
 acid linked to a promoter and a 3' non-translated sequence that
 functions in the cell to cause termination of transcription and addition
 of polyadenylated ribonucleotides to a 3' end of the mRNA molecule; and
 (2) determining a level or pattern of a molecule in a bovine cell or
 tissue comprising (a) incubating a marker nucleic acid (comprising any
 of the 15112 nucleic acid sequences or its complement or fragment) with a
 complementary nucleic acid molecule obtained from the bovine cell or
 tissue, where hybridisation between the marker nucleic acid and the
 complementary nucleic acid permits the detection of the molecule; and (b)

CC detecting the level or pattern of the complementary nucleic acid, where
 CC the detection of the complementary nucleic acid is predictive of the
 CC level or pattern of the molecule. The LMFD nucleic acid is used for
 CC determining a level or pattern of a molecule in a bovine cell or tissue.
 CC It is useful for genome mapping, gene identification and analysis, cattle
 CC breeding, preparation of constructs for use in cattle gene expression, or
 CC for genetically improving cattle. The present sequence is one of the
 CC 15112 bovine LMFD EST (expressed sequence tag) nucleic acids. Note: The
 CC present sequence was not shown in the specification put was obtained in
 CC electronic format from the USPTO web site:
 CC seqdata.uspto.gov/sequence.html?DocID=20020137139

XX Sequence 343 BP; 169 A; 56 C; 57 G; 61 T; 0 U; 0 Other;

Query Match 8.9%; Score 36; DB 7; Length 343;
 Best Local Similarity 55.6%; Pred. No. 2.4;
 Matches 69; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 36 GGGAGAGATCAAAAGAGCCCAAGAGGAAACCCAGATGAGGAAAGAAACCGCAGGA 95
 DB 71 GGGTAAAGAAAAAAGAAAAAAGAAAAAAGTCAACAAAGAGAAAAAACAACA 130
 QY 96 AATGACTACATTGAAAGAAAACTTCAAGATCAAGATAGAAAAAGCCAGAGTTTCATC 155
 DB 131 AAAGAAAAAAGAAAAAAGAAAAAAGGAAGTAAATAATCAAAAAAAGGGGGCCACC 190
 QY 156 CACT 159
 DB 191 CAAT 194

RESULT 12

ABL78974/c
 ID ABL78974 standard; cDNA; 381 BP.
 XX AC ABL78974;
 XX DT 17-MAY-2002 (first entry)
 XX DE Human ovarian cancer related cDNA clone SEQ ID NO:1952.
 XX KW Human; ovarian cancer; ovarian tumour; cytostatic; gene; ss.

XX Homo sapiens.

XX WO200192581-A2.

XX 06-DEC-2001.

XX 29-MAY-2001; 2001WO-US017756.

XX 26-MAY-2000; 2000US-0207484P.

XX (CORI-) CORIXA CORP.

XX Algate PA, Harlocker SL, Jones R;

XX WPI; 2002-122075/16.

XX Composition for therapy and diagnosis of ovarian cancer comprising
 XX polypeptide of a ovarian tumor polypeptide, polynucleotide encoding
 XX polypeptide, antibody specific to polypeptide or T cell expressing
 XX polypeptide.

XX Claim 1; SEQ ID NO 1952; 489pp; English.

XX The present invention describes a composition (I) comprising: carriers
 XX and immunostimulants; and a polypeptide (II) of a ovarian tumour
 XX polypeptide encoded by a polynucleotide (III) having a cDNA sequence (S1)
 XX from the 10912 nucleotide sequences as given in ABL77023 to ABL87934,
 XX (III) encoding (II) having a sequence (S2), a T cell population of (II),
 XX or antigen presenting cells that express (II); (I) has cytostatic
 XX activity. An oligonucleotide (IV) that hybridises to (S1) can be used for

CC of a test compound to inhibit prostate cancer in a patient; (d) assessing
CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 317 BP; 146 A; 44 C; 30 G; 96 T; 0 U; 1 Other;
Query Match 8.8%; Score 35.8; DB 5; Length 317;
Best Local Similarity 51.2%; Pred. No. 2.6;
Matches 82; Conservative 0; Mismatches 78; Indels 0; Gaps 0;
QY 14 TCCTGCGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGAGGAAACCCAG 73
Db 13 TACTTCTAGAAATTAATTAAACGCGGGGNGAAAAAATAAAAAAAAAAAAAAAAAA 72
QY 74 ATGAGGAAAGAAACGCGCAGGAAATGACTACATTGGAAGAAACITTCAGATCAAGATA 133
Db 73 AA 132
QY 134 AGAAAGCCAGAGTTTCTATCCACTTCTATCAGGAAAA 173
Db 133 AAAAAAAAAAAAAAAAAATTAATAAAAAATTTTAAAAAAA 172

RESULT 15

ABV18573
ID ABV18573 standard; cDNA; 330 BP.
XX
AC ABV18573;
XX
DT 13-SEP-2002 (first entry)
XX
DE Human prostate expression marker cDNA 18564.
XX
KW Human; prostate cancer; cytostatic; carcinogen; pharmacodynamic marker;
KW pharmacogenomic marker; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200160860-A2.
XX
PD 23-AUG-2001.
XX
PF 20-FEB-2001; 2001WO-US005171.
XX
PR 17-FEB-2000; 2000US-0183319P.
PR 16-MAR-2000; 2000US-0189862P.
PR 25-MAY-2000; 2000US-0207454P.
PR 09-JUN-2000; 2000US-0211314P.
PR 18-JUL-2000; 2000US-0219007P.
PR 13-DEC-2000; 2000US-0255281P.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Schlegel R, Endege WO, Monahan JE;
XX
DR WPI; 2001-662795/76.
XX
PT Novel isolated nucleic acid molecule associated with cancerous state of
PT prostate cells and correlating with presence of prostate cancer, useful
PT for detecting presence of prostate cancer, stage of prostate cancer.
XX
PS Claim 1; Page 3058; 11750pp; English.
XX

CC The invention relates to an isolated nucleic acid molecule (I) comprising
CC a nucleotide sequence given in Tables 1-9 (ABV00010-ABV62213) of the
CC specification or its complement. (I) is useful for: (a) assessing whether
CC a patient is afflicted with prostate cancer; (b) monitoring the
CC progression of prostate cancer in a patient; (c) assessing the efficacy
CC of a test compound to inhibit prostate cancer in a patient; (d) assessing

CC the efficacy of a therapy for inhibiting prostate cancer in a patient;
CC (e) selecting a composition for inhibiting prostate cancer in a patient;
CC (f) assessing the prostate cell carcinogenic potential of a compound; (g)
CC determining whether prostate cancer has metastasized in a patient; (h)
CC assessing the aggressiveness or indolence of prostate cancer in a patient
CC ; (i) is also useful as a pharmacodynamic or pharmacogenomic marker
XX

SQ Sequence 330 BP; 174 A; 47 C; 28 G; 57 T; 0 U; 24 Other;
Query Match 8.8%; Score 35.6; DB 5; Length 330;
Best Local Similarity 51.0%; Pred. No. 3.1;
Matches 77; Conservative 0; Mismatches 74; Indels 0; Gaps 0;
QY 8 ATTATCTCTCGGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAGCCCAAGAGGAA 67
Db 132 ATTATCTCTTTTAAATACAAATGTCNAAAAAAAAAAAAAAAAAAAAAAAAAAAA 191
QY 68 ACCCAGATGAGGAAAGAAACGCGCAGGAAATGACTACATTTGAAAGAAACTTCAAGATC 127
Db 192 AA 251
QY 128 AAGATPAGAAAGCCCAAGAGTTTTCATCCAC 158
Db 252 AAAAAAAAAAAAAAAAAANNNNGCCCCCCC 282

Search completed: April 23, 2004, 02:00:01
Job time : 295 secs

Result No.	Score	Query		Length	DB	ID	Description
		Match	%				
1	38.6	9.5	403	16	US-10-125-968-1156		Sequence 1156, Ap
2	38.4	9.5	347	13	US-10-424-599-101964		Sequence 101964, A
3	37.4	9.2	345	13	US-10-424-599-95566		Sequence 95566, A
4	36.4	9.0	325	9	US-09-764-846-30		Sequence 30, Appl
5	36.4	9.0	325	15	US-10-091-483-30		Sequence 30, Appl
6	36.2	8.9	379	13	US-10-424-599-41873		Sequence 41873, A
7	36.2	8.9	343	9	US-09-960-352-4500		Sequence 4500, Ap
8	36	8.8	381	9	US-09-867-701-1952		Sequence 1952, Ap
9	35.6	8.8	193	9	US-09-963-373-418		Sequence 418, App
10	35.6	8.8	282	13	US-10-424-599-98076		Sequence 98076, A
11	35.6	8.8	340	13	US-10-424-599-94516		Sequence 94516, A
12	35.4	8.7	332	9	US-09-960-352-11934		Sequence 11934, A
13	35.4	8.7	341	9	US-09-960-352-12302		Sequence 12302, A
14	35.2	8.7	368	13	US-10-424-599-55157		Sequence 55157, A

Db 209 AAACTCCATNTCATNTNAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAATATATAGCA 268
Qy 81 AAGAAACGGCAGAAATGACTCACTTTGAAAGAAATCTTCAAGATCAAGATTAAGAAAAG 140
Db 269 TCATAATCAAAAGGCGCCCTAGAAATAATGAAAAAGTTTCACAGAAAAAATAAAAAATG 328
Qy 141 CCAAGAAGTTTCATCCAC 158
Db 329 CAAATATCCCTTAAACAC 346

RESULT 2
US-10-424-599-101964
; Sequence 101964, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 101964
; LENGTH: 347
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_6308C.1
US-10-424-599-101964

Query Match 9.5%; Score 38.4; DB 13; Length 347;
Best Local Similarity 52.5%; Pred. No. 0.69;
Matches 84; Conservative 0; Mismatches 76; Indels 0; Gaps 0;

Qy 41 AGAATCAAAAGAGCCCAAGAAAGAAACCCAGATGAGGAAGAAAGAAACGGCAGGAATGA 100
Db 95 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 154
Qy 101 CTACATTTCAAGAAAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 160
Db 155 ATAAATAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAG 214
Qy 161 CTAATCAGAAACGAGAAATGCGAGTGGTCTTGAAGAAGT 200
Db 215 AAAATAAAAAAGAAAAAAGGGAATCCTATAAAGAAGT 254

RESULT 3
US-10-424-599-95566
; Sequence 95566, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 95566
; LENGTH: 345
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57307C.1
US-10-424-599-95566

Db 209 AAACTCCATNTCATNTNAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAATATATAGCA 268
Qy 81 AAGAAACGGCAGAAATGACTCACTTTGAAAGAAATCTTCAAGATCAAGATTAAGAAAAG 140
Db 269 TCATAATCAAAAGGCGCCCTAGAAATAATGAAAAAGTTTCACAGAAAAAATAAAAAATG 328
Qy 141 CCAAGAAGTTTCATCCAC 158
Db 329 CAAATATCCCTTAAACAC 346

Query Match 9.2%; Score 37.4; DB 13; Length 345;
Best Local Similarity 55.9%; Pred. No. 1.3; Mismatches 56; Indels 0; Gaps 0;
Matches 71; Conservative 0; Mismatches 56; Indels 0; Gaps 0;

Qy 39 AGAGATCAAAAGAGCCCAAGAAAGAAACCCAGATCAGGAAAGAAACCGCAGGAAT 98
Db 32 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 91
Qy 99 GACTACATTTGAAGAAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGAT 158
Db 92 AAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAAATAA 151
Qy 159 TTCTAAT 165
Db 152 TTGAATT 158

RESULT 4
US-09-764-846-30
; Sequence 30, Application US/09764846
; Patent No. US20020102638A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ12
; CURRENT APPLICATION NUMBER: US/09/764,846
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 348
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-764-846-30

Query Match 9.0%; Score 36.4; DB 9; Length 325;
Best Local Similarity 53.5%; Pred. No. 2.5;
Matches 76; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

Qy 6 AAATTATCTCTGCGAAAACTCAGTTGCTGGGAGAGATCAAAAGAGAGCCCAAGAGG 55
Db 182 AGATTATCTATTGCAAAAAAGATATTTCAAACCTTAATAAAAAAAGAAAAAAGAAAA 241
Qy 66 AACCAGATGAGGAAGAAAGAAACGGCAGGAATGACTACATTTGAAAGAAAACTTCAAG 125
Db 242 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 301
Qy 126 TCAAGATAAGAAAGCCAGAA 147
Db 302 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAA 323

RESULT 5
US-10-091-483-30
; Sequence 30, Application US/10091483
; Publication No. US20030049650A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PTZ12C1
; CURRENT APPLICATION NUMBER: US/10/091,483
; CURRENT FILING DATE: 2002-03-07
; NUMBER OF SEQ ID NOS: 348
; Prior Application removed - See File Wrapper or Palm
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 30
; LENGTH: 325
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-091-483-30

Sat Apr 24 08:29:36 2004

```
Query Match      9.0%; Score 36.4; DB 15; Length 325;
Best Local Similarity 53.5%; Pred. No. 2.5; Mismatches 0; Gaps 0;
Matches 76; Conservative 0; Indels 66; Indels 0; Gaps 0;

QY 6 AAATATCTCTCGGAAACTCAGTTCGCTGGGAGAGATCAAAAGAGCCCAAGAAAGG 65
Db 182 AGATTATCTATTGCAAAAGATATTTCAAACTTAAATAAAAAAAAAAAAAAAAAA 241
QY 66 AAACCCAGATCAGGAAGAAACCGCAGGAAATGACTTTCAGAGAAACTTCAAGA 125
Db 242 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 301
QY 126 TCAAGATAAGAAAGCCCAAGAA 147
Db 302 AAAAAAAAAAAAAAAAAAAAAA 323

RESULT 6
US-10-424-599-41873
; Sequence 41873, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 41873
; LENGTH: 379
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_137813C.1
US-10-424-599-41873

Query Match      8.9%; Score 36.2; DB 13; Length 379;
Best Local Similarity 53.1%; Pred. No. 3.1; Mismatches 68; Indels 0; Gaps 0;
Matches 77; Conservative 0; Indels 68; Indels 0; Gaps 0;

QY 39 AGAGATCTAAAGAGCCCAAGAAAGGAAACCCAGATCAGGAAAGAAACCGCAGGAAT 98
Db 53 AAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA 112
QY 99 GACTACATTTGAAAGAAACTTCAAGATCAAGATTAAGAAAGCCCAAGAAAGTTTCATCCAC 158
Db 113 AAAAAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 172
QY 159 TTCTAATCAGGAAACGAGAAATGGC 183
Db 173 AAAAAAAAAAAAAAAAAAAGAGGGGGGC 197

RESULT 7
US-09-960-352-4500
; Sequence 4500, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Nengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; TITLE OF INVENTION: MUSCLE AND FAT DEPOSITION
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960,352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 4500
```

```
; LENGTH: 343
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 20-BOWMS1-017-Q1-E1-E11
US-09-960-352-4500

Query Match      8.9%; Score 36; DB 9; Length 343;
Best Local Similarity 55.6%; Pred. No. 3.4; Mismatches 55; Indels 0; Gaps 0;
Matches 69; Conservative 0; Indels 55; Indels 0; Gaps 0;

QY 36 GCGAGAGATCAAAAGAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAACGCGCAGA 95
Db 71 GCGATATAAAAAAAAAAGGAAAGAAAGTCTCAAAAAGAAAGAAAGAAACAAACAA 130
QY 96 AATGACTACATTTGAAGAAAACTTCAAGATCAAGATCAAGATCAAGAAAGCCCAAGAGTTTCATC 155
Db 131 AAAGAAAAAGAAAAAGGAAAGGAAAGTAAAAAATCAAAAAAAAAAAAAAGGGGGGCCACC 190
QY 156 CACT 159
Db 191 CAAT 194

RESULT 8
US-09-867-701-1952/C
; Sequence 1952, Application US/09867701
; Patent No. US20020132237A1
; GENERAL INFORMATION:
; APPLICANT: Aglate, Paul A.
; APPLICANT: Jones, Robert
; APPLICANT: Harlocker, Susan L.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.497
; CURRENT APPLICATION NUMBER: US/09/867,701
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 10912
; SOFTWARE: Fast-Seq for Windows Version 4.0
; SEQ ID NO 1952
; LENGTH: 381
; TYPE: DNA
; ORGANISM: Homo sapien
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)-(381)
; OTHER INFORMATION: n = A,T,C or G
US-09-867-701-1952

Query Match      8.9%; Score 36; DB 9; Length 381;
Best Local Similarity 62.0%; Pred. No. 3.6; Mismatches 35; Indels 0; Gaps 0;
Matches 57; Conservative 0; Indels 35; Indels 0; Gaps 0;

QY 95 AAATGACTACATTTGAAAGAAAACTTCAAGATCAAGATCAAGATCAAGAAAGCCCAAGAGTTTCAT 154
Db 156 AATTAACCTCAGTGGAAATATTTTCAGAAAGAAAGAAAGAAATAGACAAAGCTGACTAAA 97
QY 155 CCACTTCTAATCAGGAAACGAGAAATGGCAGT 186
Db 96 CAAATATGATTATGAAATATAGGTAGGCAGT 65

RESULT 9
US-09-969-373-418
; Sequence 418, Application US/09969373
; Patent No. US2002013852A1
; GENERAL INFORMATION:
; APPLICANT: Effertz, Roger J.
; APPLICANT: Hauge, Brian M.
; TITLE OF INVENTION: Soybean SSRs and Methods of Genotyping
; FILE REFERENCE: 38-10(52679)A
; CURRENT APPLICATION NUMBER: US/09/969,373
; CURRENT FILING DATE: 2001-10-02
; PRIOR APPLICATION NUMBER: US 09/754,853
```


; Sequence 12302, Application US/09960352
; Patent No. US20020137139A1
; GENERAL INFORMATION:
; APPLICANT: Warren, Wesley C.
; APPLICANT: Tao, Mengbing
; APPLICANT: Byatt, John C.
; APPLICANT: Mathialagan, Nagappan
; TITLE OF INVENTION: NUCLEIC ACID AND OTHER MOLECULES ASSOCIATED WITH LACTATION AND
; FILE REFERENCE: 16511.006/37-21(10298)C
; CURRENT APPLICATION NUMBER: US/09/960.352
; CURRENT FILING DATE: 2001-09-24
; NUMBER OF SEQ ID NOS: 15112
; SEQ ID NO 12302
; LENGTH: 341
; TYPE: DNA
; ORGANISM: Bos taurus
; OTHER INFORMATION: Clone ID: 53-BOWMS1-009-Q1-E1-F2
US-09-960-352-12302

Query Match 8.7%; Score 35.4; DB 9; Length 341;
Best Local Similarity 53.2%; Pred. No. 5.1;
Matches 75; Conservative 0; Mismatches 66; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAGAAACCCAGATGAGGAAAGAAACCGGCGAGGAAAT 98
DB 267 AAAAAAGAAAAACAAAAAGAAAAAACAACAAAAAGAAAAAAGAAAAA 208

QY 99 GACTACATTTGAAGAAACTTCAAGATCAAGTAAGAAAGCCCAAGAGTTTCATCCAC 158
DB 207 AAAAAAAGAAAAAAGAAAGACAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 148

QY 159 TTCTAATCAGGAAACGAGAA 179
DB 147 AAAAAAAGAAAAAAGAAAAA 127

RESULT 14
US-10-424-599-55157
; Sequence 55157, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 55157
; LENGTH: 368
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(368)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_20818C.1
US-10-424-599-55157

Query Match 8.7%; Score 35.2; DB 13; Length 368;
Best Local Similarity 51.2%; Pred. No. 6;
Matches 82; Conservative 0; Mismatches 78; Indels 0; Gaps 0;

QY 39 AGAGATCAAAAGAGCCCAAGAGAAACCCAGATGAGGAAAGAAACCGGCGAGGAAAT 98
DB 9 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 68

QY 99 GACTACATTTGAAGAAACTTCAAGATCAAGTAAGAAAGCCCAAGAGTTTCATCCAC 158

Db 59 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 128

QY 159 TTCTAATCAGGAAACGAGAAATGGCAGTGGTTCGAAGAA 198
DB 129 TTAAAGAAATCAAGTTTACTTCCCGGTCTTCCAAGAA 168

RESULT 15
US-10-424-599-89471
; Sequence 89471, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 89471
; LENGTH: 327
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(327)
; OTHER INFORMATION: unsure at all n locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_51802C.1
US-10-424-599-89471

Query Match 8.6%; Score 35; DB 13; Length 327;
Best Local Similarity 59.0%; Pred. No. 6.5;
Matches 59; Conservative 0; Mismatches 41; Indels 0; Gaps 0;

QY 18 GCGAAACTCAGTTGCTGGGAGAGAAATCAAAAGAAAGCCCAAGAAAGGAAACCCAGATCA 77
DB 209 GCGAAACCCACACGGGAAGCCCGCACAGAAACCCGACACGACGAAAGAAACGAGAGGC 268

QY 78 GGAAGAAACGCGCAGGAAATGACTACATTTGAAAGAAAA 117
DB 269 GGAAGAAACAGCCCGCAAAACAAACAGCGGCGAGGCAAA 308

Search completed: April 23, 2004, 04:03:59
Job time : 298 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: April 22, 2004, 18:37:20 ; Search time 64 Seconds
(without alignments)
3511.799 Million cell updates/sec

Title: US-10-005-907-1_COPY_25_429
Perfect score: 405
Sequence: 1 atgggaattatctctctg.....atgaattgtgttccacac 405

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1083176

Minimum DB seq length: 0
Maximum DB seq length: 405

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents NA.*
1: /cgn2_6/ptodata/2/ina/5A.COMB.seq.*
2: /cgn2_6/ptodata/2/ina/5B.COMB.seq.*
3: /cgn2_6/ptodata/2/ina/6A.COMB.seq.*
4: /cgn2_6/ptodata/2/ina/6B.COMB.seq.*
5: /cgn2_6/ptodata/2/ina/PCTUS.COMB.seq.*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	34.8	8.6	399	4	US-09-621-976-8976
2	33.6	8.3	249	4	US-09-621-976-1322
3	33.4	8.2	240	4	US-09-621-976-1324
4	33	8.1	400	4	US-08-956-171E-4053
5	31.8	7.9	400	4	US-08-956-171E-3835
6	31.6	7.8	333	3	US-09-018-584A-27
7	31	7.7	240	1	US-08-628-417-6
8	31	7.7	372	3	US-09-018-584A-13
9	30.8	7.6	260	2	US-08-520-678A-29
10	30.8	7.6	260	3	US-08-897-126-29
11	30.8	7.6	377	4	US-08-956-171E-4397
12	30.8	7.6	396	4	US-09-640-173-16
13	30.8	7.6	396	4	US-09-713-550-16
14	30.4	7.5	356	2	US-08-520-678A-22
15	30.4	7.5	356	3	US-08-897-126-22
16	30.2	7.5	144	1	US-08-702-344-26
17	30.2	7.5	147	4	US-09-621-976-10383
18	30	7.4	291	1	US-07-922-723A-7
19	30	7.4	291	1	US-07-799-828C-7
20	30	7.4	291	1	US-08-074-275-7
21	30	7.4	291	1	US-08-480-366-7
22	30	7.4	291	2	US-07-952-277A-7
23	29.8	7.4	251	4	US-08-956-171E-4555
24	29.8	7.4	253	2	US-08-520-678A-25
25	29.8	7.4	253	3	US-08-897-126-25
26	29.6	7.3	248	3	US-09-007-005-32
27	29.6	7.3	248	3	US-09-244-796-32

28	29.6	7.3	277	3	US-09-007-005-3	Sequence 3, Appli
29	29.6	7.3	277	3	US-09-244-796-3	Sequence 3, Appli
30	29.6	7.3	373	2	US-08-557-309B-4	Sequence 4, Appli
31	29.6	7.3	373	3	US-08-834-306-4	Sequence 4, Appli
32	29.6	7.3	373	3	US-08-993-674A-4	Sequence 4, Appli
33	29.6	7.3	373	4	US-09-256-976-4	Sequence 20, Appli
34	29.4	7.3	121	3	US-09-297-535-20	Sequence 12892, A
35	29.4	7.3	130	4	US-09-621-976-12892	Sequence 10254, A
36	29.4	7.3	147	4	US-09-621-976-10254	Sequence 16692, A
37	29.4	7.3	164	4	US-08-910-632-4	Sequence 4, Appli
38	29.4	7.3	204	3	US-08-805-631A-4	Sequence 4, Appli
39	29.4	7.3	204	4	US-09-569-344-4	Sequence 30, Appli
40	29.4	7.3	270	2	US-08-520-678A-30	Sequence 30, Appli
41	29.4	7.3	270	3	US-08-897-126-30	Sequence 20, Appli
42	29.4	7.3	321	3	US-09-018-584A-20	Sequence 18, Appli
43	29.4	7.3	376	2	US-08-623-906A-18	Sequence 5, Appli
44	29.4	7.3	383	3	US-09-155-942-5	
45	29.4	7.3				

ALIGNMENTS

RESULT 1

US-09-621-976-8976
; Sequence 8976, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 8976
; LENGTH: 399
; TYPE: DNA
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-621-976-8976

Query Match	8.6%	Score 34.8;	DB 4;	Length 399;
Best Local Similarity	12.4%	Pred. No. 0.26;		
Matches	30;	Conservative 110;	Mismatches 102;	Indels 0; Gaps 0;
QY	3	GGGAAATTATCTCTCGGAAACTCAGTTGCTGGAGAGATCAAAAGAACCCAGAA	62	
Db	42	KKRRRRRAWWMMKSCMCKSKSWRSWGMVKMKGRGAASWAGYMSWMTYNTMRWY	101	
QY	63	AGGAAACCAGATGAGGAGAAAGAACGGCAGAGAAATGACTATTGAAAGAAACTCA	122	
Db	102	YRKYKACTKWRAGWNGKAWWAWAYAKWYMAWERTAMKWMKSKSRMRRAWYA	161	
QY	123	AGATCAAGTAAGAAAGCCAGAGTTTCATCCACTTCTAATCAGGAAACAGAAATCG	182	
Db	162	WMTYMAARTTWGMRASCYRGAYMASAGMYWMTYMRKMWYSAGWSMKRKTWRCS	221	
QY	183	CAGTGTCTTGAGAGTGTGCTACACTGCTAATCATCCATCCCATCAGAGATCCTC	242	
Db	222	YSCWSYCMWGAOMWMTKTSRWSYWSYISRCTKYRSRCCMSYWKTYRSWYCASC	281	
QY	243	CC 244		
Db	282	YY 283		

RESULT 2

US-09-621-976-1322
; Sequence 1322, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:

NUMBER OF SEQUENCES: 5256
CORRESPONDENCE ADDRESS:
ADDRESSEE: Human Genome Sciences, Inc.
STREET: 9410 Key West Avenue
CITY: Rockville
STATE: Maryland
COUNTRY: USA
ZIP: 20850

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette, 3.50 inch, 1.44Mb storage
COMPUTER: HP Vectra 486/33
OPERATING SYSTEM: MS-DOS version 6.2
SOFTWARE: ASCII Text

CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/956,171E
FILING DATE: 20-Oct-1997
CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:
APPLICATION NUMBER: 60/009,861
FILING DATE: January 5, 1996
APPLICATION NUMBER: 08/781,986
FILING DATE: January 3, 1997

ATTORNEY/AGENT INFORMATION:
NAME: Mark J. Hyman
REGISTRATION NUMBER: 46,789
REFERENCE/DOCKET NUMBER: P3248P1
TELECOMMUNICATION INFORMATION:
TELEPHONE: (240) 314-1224
TELEFAX: (301) 309-8439

INFORMATION FOR SEQ ID NO: 3835:
SEQUENCE CHARACTERISTICS:
LENGTH: 400 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: linear
SEQUENCE DESCRIPTION: SEQ ID NO: 3835:

US-08-956-171E-3835

Query Match 7.9%; Score 31.8; DB 4; Length 400;
Best Local Similarity 51.8%; Pred. No. 2.2; Indels 0; Gaps 0;
Matches 72; Conservative 0; Mismatches 67;

QY 41 AGAATCAAAAGAGCCCAAGAGAAAGGAAAGCCAGATGAGGAAAGAAAGGAAAGGAAAGTGA 100
DB 80 ATAATAATCTTATTACACAGAGATATATTACTTATGCAAGGGCGGAGGAAATCATCATGT 139
QY 101 CTACATTTCAAGAAAGAACTTCAAGATCAGATAGAGAAAGCCCAAGAGTTTCATCCACTT 160
DB 140 CTATTACTGAAAGAAACAGCTCAGCAAGCTGATTAATTAATTAATTAATTAATTAATTAATG 199
QY 161 CTAATCAGGAAAGCAGAA 179
DB 200 CGAATGATTAAGAGGAA 218

RESULT 6
US-09-018-584A-27
Sequence 27, Application US/09018584A
Patent No. 6238863
GENERAL INFORMATION:
APPLICANT: Schumm, James W.
APPLICANT: Bacher, Jeffery W.
TITLE OF INVENTION: MATERIALS AND METHODS FOR
IDENTIFYING AND ANALYZING INTERMEDIATE TANDEM
REPEAT DNA MARKERS
NUMBER OF SEQUENCES: 147
CORRESPONDENCE ADDRESS:
ADDRESSEE: Promega Corporation
STREET: 2800 Woods Hollow Road
CITY: Madison
STATE: Wisconsin
COUNTRY: U.S.A.
ZIP: 53711-5399

COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette - 3.5 inch, 1.44 Mb
COMPUTER: IBM compatible PC
OPERATING SYSTEM: Windows 95
SOFTWARE: Word 97 (DOS text format)
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/018,584A
FILING DATE: 04-Feb-1998
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Grady J. Frenchick
REGISTRATION NUMBER: 29,018
REFERENCE/DOCKET NUMBER: 16026.9180
TELECOMMUNICATION INFORMATION:
TELEPHONE: (608) 257-3501
TELEFAX: (608) 257-2275
INFORMATION FOR SEQ ID NO: 27:
SEQUENCE CHARACTERISTICS:
LENGTH: 333 bp
TYPE: Nucleic Acid
STRANDEDNESS: Double
TOPOLOGY: Circular
MOLECULE TYPE: Genomic DNA
HYPOTHETICAL: no
IMMEDIATE SOURCE:
LIBRARY: plasmid, pGem3zf(+)
CLONE: G539
POSITION IN GENOME:
CHROMOSOME/SEGMENT: 15q26.2

US-09-018-584A-27

Query Match 7.8%; Score 31.6; DB 3; Length 333;
Best Local Similarity 56.9%; Pred. No. 2.3; Indels 0; Gaps 0;
Matches 58; Conservative 0; Mismatches 44;

QY 38 GAGAGATCAAAAGAGCCCAAGAGAAAGGAAAGCCAGATGAGGAAAGAAAGGAAAGGAAAGGAAAG 97
DB 78 GAAAG 137
QY 98 TGACTACATTTGAAAGAAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 139
DB 138 GAAAG 179

RESULT 7
US-08-628-417-6
Sequence 6, Application US/08628417
Patent No. 5627054
GENERAL INFORMATION:
APPLICANT: GILLESPIE, DAVID
TITLE OF INVENTION: COMPETITOR PRIMER ASYMMETRIC
POLYMERASE CHAIN REACTION
NUMBER OF SEQUENCES: 7
CORRESPONDENCE ADDRESS:
ADDRESSEE: U.S. ARMY CHEMICAL AND BIOLOGICAL
DEFENSE COMMAND
STREET: OFFICE OF THE CHIEF COUNSEL (AMSCB-GC)
CITY: ABERDEEN PROVING GROUND
STATE: MARYLAND
COUNTRY: USA
ZIP: 21010-5423
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/628,417
FILING DATE:
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: BIFFONI, ULYSSES J
REGISTRATION NUMBER: 39,908

;; TITLE OF INVENTION: METHODS OF USE THEREFOR
;; FILE REFERENCE: 210121.484C2
;; CURRENT APPLICATION NUMBER: US/09/640.173
;; CURRENT FILING DATE: 2000-08-15
;; NUMBER OF SEQ ID NOS: 196
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 16
;; LENGTH: 396
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)-(396)
;; OTHER INFORMATION: n = A,T,C or G
US-09-640-173-16

Query Match 7.6%; Score 30.8; DB 4; Length 396;
Best Local Similarity 51.2%; Pred. No. 4.4;
Matches 65; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 21 AAAAAGTCAAGTTCCTGGGAGAGATCAAAAGAGCCCAAGAGGAAAGCCAGATGAGGA 80
DB 131 AAAATTTNNCCCCNAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 72
QY 81 AAGAAACCGCAGGAAATGACTTACATTTTGAAGAAATCTCAAGATCAAGTAAAGAAAG 140
DB 71 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 12
QY 141 CCAAGAA 147
DB 11 AAAAAAA 5

RESULT 13
US-09-713-550-16/c
;; Sequence 16, Application US/09713550
;; Patent No. 6617109
;; GENERAL INFORMATION:
;; APPLICANT: Xu, Jiangchun
;; APPLICANT: Stolk, John A.
;; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
;; TREATMENT OF OVARIAN CANCER
;; TITLE OF INVENTION: THERAPY AND DIAGNOSIS OF OVARIAN CANCER
;; FILE REFERENCE: 210121.484C4
;; CURRENT APPLICATION NUMBER: US/09/713,550
;; CURRENT FILING DATE: 2000-11-14
;; NUMBER OF SEQ ID NOS: 205
;; SOFTWARE: FastSeq for Windows Version 3.0
;; SEQ ID NO 16
;; LENGTH: 396
;; TYPE: DNA
;; ORGANISM: Homo sapien
;; FEATURE:
;; NAME/KEY: misc feature
;; LOCATION: (1)-(396)
;; OTHER INFORMATION: n = A,T,C or G
US-09-713-550-16

Query Match 7.6%; Score 30.8; DB 4; Length 396;
Best Local Similarity 51.2%; Pred. No. 4.4;
Matches 65; Conservative 0; Mismatches 62; Indels 0; Gaps 0;
QY 21 AAAAAGTCAAGTTCCTGGGAGAGATCAAAAGAGCCCAAGAGGAAAGCCAGATGAGGA 80
DB 131 AAAATTTNNCCCCNAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 72
QY 81 AAGAAACCGCAGGAAATGACTTACATTTTGAAGAAATCTCAAGATCAAGTAAAGAAAG 140
DB 71 AAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAAAGAAAAA 12
QY 141 CCAAGAA 147
DB 11 AAAAAAA 5

RESULT 14
US-08-520-678A-22/c
;; Sequence 22, Application US/08520678A
;; Patent No. 5874565
;; GENERAL INFORMATION:
;; APPLICANT: Rice, Charles M.
;; APPLICANT: Kolykhalov, Alexander A.
;; TITLE OF INVENTION: NOVEL 3' TERMINAL SEQUENCE OF HEPATITIS
;; TITLE OF INVENTION: C VIRUS GENOME AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF
;; NUMBER OF SEQUENCES: 39
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Howell & Haferkamp, L.C.
;; STREET: 7733 Forsyth Blvd., Suite 1400
;; CITY: St. Louis
;; STATE: MO
;; COUNTRY: USA
;; ZIP: 63105
;; COMPUTER READABLE FORM:
;; MEDIUM TYPE: Floppy disk
;; COMPUTER: IBM PC compatible
;; OPERATING SYSTEM: PC-DOS/MS-DOS
;; SOFTWARE: PatentIn Release #1.0, Version #1.30
;; CURRENT APPLICATION DATA:
;; APPLICATION NUMBER: US/08/520,678A
;; FILING DATE:
;; CLASSIFICATION: 536
;; ATTORNEY/AGENT INFORMATION:
;; NAME: Henderson, Melodie W.
;; REGISTRATION NUMBER: 37,848
;; REFERENCE/DOCKET NUMBER: 6029-6836
;; TELECOMMUNICATION INFORMATION:
;; TELEPHONE: 314-727-5188
;; TELEFAX: 314-727-6092
;; TELEX:
;; INFORMATION FOR SEQ ID NO: 22:
;; SEQUENCE CHARACTERISTICS:
;; LENGTH: 356 base pairs
;; TYPE: nucleic acid
;; STRANDEDNESS: single
;; TOPOLOGY: linear
;; MOLECULE TYPE: DNA (genomic)
US-08-520-678A-22
Query Match 7.5%; Score 30.4; DB 2; Length 356;
Best Local Similarity 48.8%; Pred. No. 5.6; Indels 0; Gaps 0;
Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 48 AAAGAAGCCCAAGAAAGGAAACCCAGATGAGGAGAAAGAAAGCCAGAGAAATGACTACATT 107
DB 251 AAGGAAGGAAAGAAAGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 192
QY 108 TGAAGAAAGAACTTCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCAAGATCA 167
DB 191 AAAAAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 132
QY 168 GCAAAACGAGATGCGAGTGGTTCTTGAAGAAGTGTCTACACTGTCTAT 215
DB 131 AAACAGGAATGCGCTTAAGAGCGCGAGTGTCTTACCCCAACCTTCAT 84
RESULT 15
US-08-897-126-22/c
;; Sequence 22, Application US/08897126
;; Patent No. 6297003
;; GENERAL INFORMATION:
;; APPLICANT: Rice, Charles M.
;; APPLICANT: Kolykhalov, Alexander A.
;; TITLE OF INVENTION: NOVEL 3' TERMINAL SEQUENCE OF HEPATITIS
;; TITLE OF INVENTION: C VIRUS GENOME AND DIAGNOSTIC AND THERAPEUTIC USES THEREOF
;; NUMBER OF SEQUENCES: 39
;; CORRESPONDENCE ADDRESS:
;; ADDRESSEE: Howell & Haferkamp, L.C.

STREET: 7733 Forsyth Blvd., Suite 1400
CITY: St. Louis
STATE: MO
COUNTRY: USA
ZIP: 63105
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/897,126
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/520,678
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Henderson, Melodie W.
REGISTRATION NUMBER: 37,848
REFERENCE/DOCKET NUMBER: 6029-6836
TELECOMMUNICATION INFORMATION:
TELEPHONE: 314-727-5188
TELEFAX: 314-727-6092
TELEX:
INFORMATION FOR SEQ ID NO: 22:
SEQUENCE CHARACTERISTICS:
LENGTH: 356 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
US-08-897-126-22

Query Match
Best Local Similarity 48.8%; Score 30.4; DB 3; Length 356;
Matches 82; Conservative 0; Mismatches 86; Indels 0; Gaps 0;
QY 48 AAAGAAGCCCAAGAAAGGAAACCCAGATGAGGAAAGAAAACGGCAGGAAATGACTACATT 107
Db 251 AAGGAGGAAAGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAGGAAAG 192
QY 108 TGAAGAAACCTTCAAGATCAAGATAAGAAAGCCAAAGTTTCATCCACTTCTAATCA 167
Db 191 AAAAAAAAAAGAAA 132
QY 168 GGAACACGAGATGGCAGTGGTTCTCAAGAACTGTGCTACACTGTCAT 215
Db 131 AAACACGGAATGGCCTAAGAGCGCGAGTGTATTACCCCACTTCAT 84

Search completed: April 23, 2004, 03:09:53
Job time : 65 secs

This Page Blank (uspto)

This Page Blank (uspto)

This Page Blank (uspto)